

# Systematic effects on the size-luminosity relation: dependence on model fitting and morphology

M. Bernardi<sup>1\*</sup>, A. Meert<sup>1</sup>, V. Vikram<sup>1</sup>, M. Huertas-Company<sup>2</sup>, S. Mei<sup>2</sup>, F. Shankar<sup>2</sup>  
& R. K. Sheth<sup>1,3</sup>

<sup>1</sup>*Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA*

<sup>2</sup>*GEPI, Observatoire de Paris, CNRS, Univ. Paris Diderot; Place Jules Janssen, 92190 Meudon, France*

<sup>3</sup>*The Abdus Salam International Center for Theoretical Physics, Strada Costiera 11, 34151 Trieste, Italy*

Accepted . Received ; in original form

## ABSTRACT

We quantify the systematics in the size-luminosity relation of galaxies in the SDSS main sample (i.e. at  $z \sim 0.1$ ) which arise from fitting different one- and two-component model profiles to the images. We use a novel argument to show that photometric information alone indicates that the majority of galaxies are not single-component systems. A SerExp model, a Sersic bulge plus exponential disk, for the two-components, is a better description than is a single Sersic profile; it is also better than the traditional deVaucouleurs bulge plus exponential disk model. In objects brighter than  $L_*$ , fitting a single Sersic profile to what is really a two-component SerExp system leads to biases: the half-light radius is increasingly overestimated as  $n$  of the fitted single component increases; it is also overestimated at  $B/T \sim 0.6$ . For such objects, the assumption of a single Sersic component is particularly misleading. However, the net effect on the size-luminosity relation is small, except for the most luminous tail.

We then study how this relation depends on morphology. Our analysis is one of the first to use Bayesian-classifier derived weights, rather than hard cuts, to define morphology. Crudely, there appear to be only two relations: one for early-types, and the other for later-types (Es, S0s and Sa's are early-types, whereas Sbs and Scds are late). Closer inspection shows that within the early-type sample S0s tend to be 15% smaller than Es of the same luminosity, and, among faint late types, Sbs are more than 25% smaller than Scds. Neither the early- nor the late-type relations are pure power-laws: both show significant curvature, which we quantify. This curvature confirms that two mass scales are special for early-type galaxies:  $M_* = 3 \times 10^{10} M_\odot$  and  $2 \times 10^{11} M_\odot$ . At  $M_* > 2 \times 10^{11} M_\odot$ , the  $R - M_*$  relation is the same as that of the bulge component. These same mass scales are also special for late types: there is almost no correlation between  $R$  and  $M_*$  below the former, and almost no late-types above the latter.

Our bulge-disk decompositions indicate that disks in both early- and late-types tend to be  $\sim 3\times$  larger than bulges. Although the  $R - L$  and  $R - M_*$  relations of disks of late types run parallel to those of the total light in late-types, these relations for the bulges of early-types show almost no curvature, and lie well below those for the total light in early-types especially at low  $L$  and  $M_*$ . In fact, for narrow bins in velocity dispersion  $\sigma$ ,  $R_{\text{bulge}} \propto L_{\text{bulge}}$  (and similarly  $R_{\text{bulge}} \propto M_{*\text{bulge}}$ ): our SerExp bulges satisfy the virial theorem scaling, with bulges of fixed  $L$  but smaller  $\sigma$  having larger sizes. The intrinsic scatter in the  $R - L$  relation decreases at large luminosity and/or stellar mass and should provide additional constraints on models of how the most massive galaxies formed.

**Key words:** galaxies: structural parameters – galaxies: fundamental parameters – galaxies: evolution

## 1 INTRODUCTION

The spatial (and color) distribution of star light in a galaxy is thought to encode information about its formation history.

\* E-mail: bernardm@sas.upenn.edu

In recent years, the correlation between size and luminosity for early-type galaxies has received much attention, because high redshift early-types appear to be more compact than their counterparts at low redshift (e.g. Trujillo et al. 2006; van Dokkum et al. 2008; Cimatti et al. 2008; Bruce et al. 2012). However, both the size and the luminosity estimates are derived parameters, obtained by fitting to the observed surface brightness distribution. As a result, they depend on assumptions about the intrinsic shape of the surface brightness profile. E.g., if the fit assumes that galaxies are made up of two components or just one, and if two, whether they are modelled as the sum of an exponential and a deVaucouleurs (1948) profile, an exponential and a Sersic (1968), or two Sersics.

The main goal of this paper is to quantify the systematics on the local  $R - L$  relation which are associated with the choice of a particular model. In practice, ‘local’ means the  $z \sim 0.1$  galaxies in DR7 of the SDSS Main Galaxy sample (Abazajian et al. 2009). Because this sample is apparent magnitude limited, in practice, by  $R - L$  relation we always mean  $\log_{10} R$  fitted as a function of absolute magnitude (see Sheth & Bernardi 2012 for a simple description of the bias which would arise from fitting  $L$  as a function of  $R$ ). And  $R$  denotes the radius which encloses half the total light  $L$ . (For exponential disks, this radius is 1.67 times the scale length of the exponential.)

Our goal implies that we must fit the observed profiles and determine the associated  $R - L$  relation using a variety of different models. We perform the fits using the **PyMorph** package, which can fit seeing convolved two components models to observed surface brightness profiles (Vikram et al. 2010). The algorithm is described and tested in Meert et al. (2012a,b). Tests on synthetic images show that when the fitted functional form is the same as the one used to generate the image, then **PyMorph** returns accurate values of the free-parameters (e.g., total light, half-light radius, bulge-total ratio).

At first, we explore results from a single Sersic profile since the standard to date has been to use parameters from single Sersic fits. We will demonstrate that fits to a single Sersic profile are likely to return biased estimates of  $R$  and  $L$ , whereas fits to the sum of an exponential and Sersic profile should be less biased. Section 2 shows that the **PyMorph** derived  $R - L$  relation for single Sersic fits is offset from the standard in the literature (Shen et al. 2003). However, it is consistent with that obtained from a more recent determination of the Sersic parameters (Simard et al. 2011; hereafter S11). Although we believe the S11 reductions are to be preferred to those in Shen et al., Appendix A shows that they too are slightly biased – they appear to imply evolution in the  $n - L$  and  $R - L$  relations which we believe to be unphysical. This evolution is not present in the **PyMorph** reductions, so the remainder of the paper considers **PyMorph** exclusively.

Section 3 compares the  $R - L$  relation based on single Sersic, Sersic + exponential, deVaucouleur + exponential, and single deVaucouleur fits, showing that the relations from single Sersic fits are offset to larger sizes and those from single deVaucouleur fits to smaller sizes, compared to those from the two-component fits. (In all cases, ‘size’ means the radius which contains half of the total light: in the case of two components, this is a complicated function of the light

in each component and the scale radii.) Appendix B argues that although fits to a single Sersic profile are likely to return biased estimates of  $R$  and  $L$ , whereas fits to the sum of an exponential and Sersic profile should be less biased, the net effect on the derived  $R - L$  relation is small. On the other hand, fitting a realistic model is necessary to obtain sensible estimates of the intrinsic scatter around the mean  $R - L$  relation. This is the subject of Section 3.3, where we show that the intrinsic scatter correlates with  $\sigma$  and decreases with increasing luminosity.

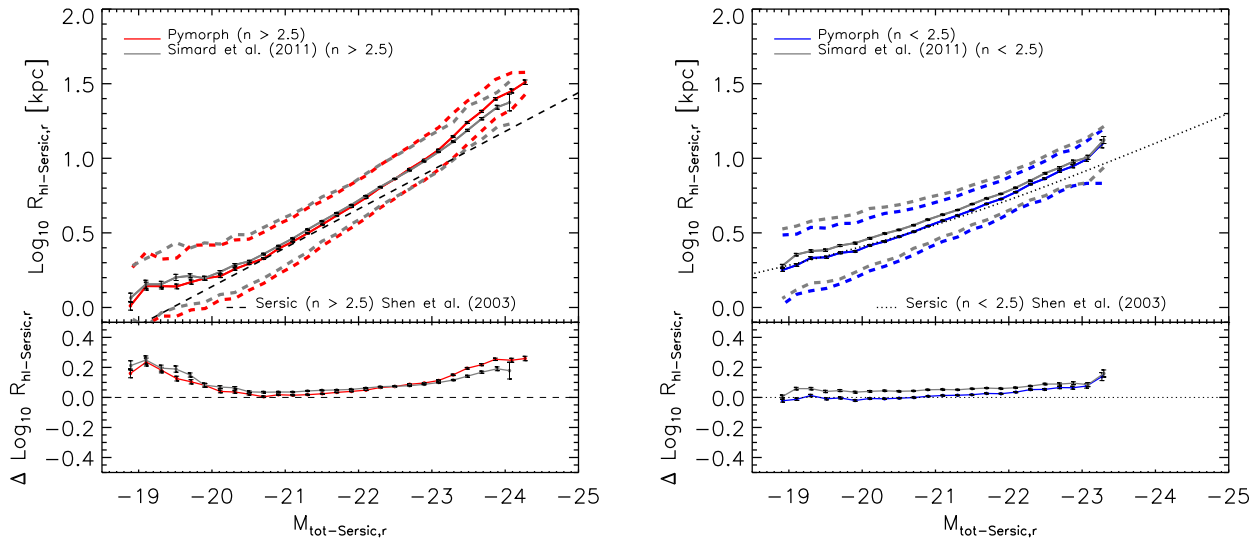
Section 3.4 studies how the  $R - L$  relation depends on morphological type, using both the eye-ball classifications of Fukugita et al. (2007; hereafter F07) and Nair et al. (2010; hereafter N10) and the Bayesian Automated Classifications (hereafter BAC) of Huertas-Company et al. (2011). The latter are particularly interesting, because they are expressed as probabilistic weights – something we expect will become increasingly common in the next generation of large datasets. We explore the use of hard cuts based on these weights as indicators of morphology, as well as simply weighting each galaxy by the probability that it is one type or another.

Bulge dominated galaxies do have disk components, and disk dominated galaxies have small bulges. In Section 4 we use our Sersic-exponential fits to study the  $R - L$  relation of the bulge components of the former and the disk components of the latter. In contrast to the curved relations we see for the total size-luminosity (size- $M_*$ ) correlation, the  $R - L$  ( $R - M_*$ ) relation for bulges of early-types appears to be a power law, the slope of which is the same as that of the early-type  $R - L$  ( $R - M_*$ ) relation at fixed velocity dispersion. In addition, it shows that the two mass scales which are important for early-type galaxies,  $M_* = 3 \times 10^{10} M_\odot$  and  $2 \times 10^{11} M_\odot$  (Bernardi et al. 2011a,b) are also special for late types. A final section summarizes our findings.

When converting from angular sizes and apparent brightnesses to physical sizes and luminosities, we assume a flat  $\Lambda$ CDM model with  $\Omega_m = 0.23$  and a Hubble constant whose present value is  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

## 2 THE SIZE-LUMINOSITY RELATION AT $Z \sim 0.1$ FROM SINGLE SERSIC FITS

There is an analytic expression for the light enclosed within a given distance of the center of a circular Sersic profile. From this, the half light radius can be obtained easily. However, if the object has axis ratio  $b/a \neq 1$ , where  $b$  and  $a$  are the half-light radii along the principal axes of the image, then the corresponding expression must be integrated numerically. Since this can be time-consuming, it is usual to approximate this case by using the expression for a circle, but with a suitably chosen effective circular radius. The most common choice is  $\sqrt{ba} = a \sqrt{b/a}$ , but Saglia et al. (2010) have recently shown that  $(b + a)/2$  is more accurate: for bulge dominated systems the difference matters little, but it does matter for disks. Therefore, in the next subsection (2.1), where we compare with previous work, we use  $\sqrt{ba}$ . Thereafter, we use  $(b + a)/2$ .



**Figure 1.** The r-band single-Sersic based half-light radius ( $R_{hl}$ ) versus total absolute magnitude ( $M_{tot}$ ) relation for objects with  $n > 2.5$  (left) and  $n < 2.5$  (right). In the panel on the left, our **PyMorph** determination is in good agreement with that based on single-Sersic parameters from Simard et al. (2011), but lies about 0.1 dex above, and is more curved than the fit reported by Shen et al. (2003). Symbols with error bars (joined by a solid curve for clarity) show the median half-light radius in bins of absolute magnitude. Dashed lines show the 16th and 84th percentile. In the panel on the right (objects with  $n < 2.5$ ), except for the brightest objects, the **PyMorph** relation lies systematically 0.05 dex below that of Simard et al.

## 2.1 Comparison with previous work

Using the objects in an earlier SDSS data release, Shen et al. (2003) reported fits to the  $R-L$  relation for objects which had  $n > 2.5$  and  $n < 2.5$ , where  $R$ ,  $L$  and  $n$  were determined from a single Sersic fit to the light profile. Note that the Sersic parameters used by Shen et al. were estimated from a 1-dimensional radial surface brightness profile (**profMean**), measured in  $\sim 5-10$  azimuthally averaged annuli (Blanton et al. 2003). Thus, it is expected to be significantly less accurate than a 2-dimensional fit to the whole galaxy image.

The Shen et al. relations for  $n > 2.5$  and  $n < 2.5$  are shown as the dashed and dotted lines in the left and right hand panels of Figure 1, respectively. The red and blue symbols with error bars show our determination of the single-Sersic based relation, where now  $R$ ,  $L$  and  $n$  are from our **PyMorph** reductions, and the grey symbols and error bars show the  $R-L$  relation which follows from single-Sersic fits performed by Simard et al. (2011, hereafter S11).

For objects with  $n < 2.5$ , the S11-derived relation runs parallel to that from Shen et al., but is offset to larger sizes by 0.05 dex, whereas the **PyMorph**-derived relation transitions from Shen et al. at low luminosities to S11 at high luminosities. For objects with  $n > 2.5$  the **PyMorph**-derived relation lies about 0.1 dex above, and is more curved than the fit reported by Shen et al. The **PyMorph** and S11 based relations depart significantly from Shen et al. at the low and high luminous ends, where they curve upwards to larger sizes. For this reason, we are inclined to conclude that, at least at the bright end, Shen et al. is slightly biased. At the low end the curvature could be due to contamination by later-type galaxies.

However, at the highest luminosities of objects with  $n > 2.5$ , the **PyMorph** and S11 relations are also slightly but significantly different from one another. Appendix A shows

that, in fact, at high luminosities, the derived magnitudes and sizes can be quite different: the correlated nature of these differences means that the  $R-L$  relation is only moderately affected. Appendix A goes on to show that the S11 reductions appear to require rather dramatic evolution in  $n$  and  $R$ : both are larger at  $z = 0.2$  than at  $z = 0.05$ . Since we believe this is unphysical, we conclude that the **PyMorph** reductions, which show no such systematic trend with  $z$ , are less biased, so we will use them in the remainder of this paper.

## 3 THE SIZE-LUMINOSITY RELATION AT $Z \sim 0.1$ FROM ONE- AND TWO-COMPONENT FITS

In this section we study how the  $R-L$  relation depends on the functional form for the surface brightness profile that was assumed when estimating  $R$  and  $L$ . We would especially like to compare the effects of fitting one versus two-component models to the images. It is conventional to speak of these as ‘bulge’ and ‘disk’ components; while this is accurate for disk-dominated systems (typically later-type galaxies), it may be better to think of the ‘disk’ component in bulge-dominated systems (typically early-type galaxies) as simply being a second component that is not necessarily a (thin, inclined) disk.

We stated earlier that, for each component, we can approximate the half-light radius by assuming the image is a circle of radius  $(b+a)/2$ . But what should we do when we have two components? A natural choice would be to circularize each component using its own  $(b+a)/2$ , and to then determine the half light radius of the sum of the circularized components, where each is weighted by the fraction of

Selection	Ell	S0	Sa	Sb	Scd
EARLY-TYPES					
Selected					
P(E+S0) > 0.85 AND $n > 3$	0.70	0.21	0.08	0.01	0
$n > 2.5$	0.44	0.18	0.20	0.13	0.05
Missed					
P(E+S0) < 0.85 OR $n < 3$	0.10	0.43			
$n < 2.5$	0.02	0.12			
LATE-TYPES					
Selected					
P(E+S0) < 0.15 AND $n < 3$	0	0.01	0.08	0.36	0.51
$n < 2.5$	0.01	0.04	0.11	0.33	0.45
Missed					
P(E+S0) > 0.15 OR $n > 3$			0.86	0.47	0.21
$n > 2.5$			0.75	0.41	0.16

**Table 1.** Eyeball morphological classifications from Fukugita et al. (2007). We set Ell ( $T = 0$  and  $0.5$ ), S0 ( $T = 1$ ), Sa ( $T = 1.5$  and  $2$ ), Sb ( $T = 2.5$  and  $3$ ), and Scd ( $T = 3.5, 4, 4.5, 5$ , and  $5.5$ ).

the total light that it contains (e.g. equation C3). We have found that this approximation is quite accurate.

### 3.1 Selection of morphological types based on hard-cuts

Since the systematic biases may differ for bulge or disk dominated systems, we would like to separate out the effects of morphology from those on the functional form. Therefore, we will begin by first selecting objects of a single morphological type.

In practice this is difficult, because unambiguous determinations of the morphological type are not straightforward, although the task is slightly easier for bulge dominated systems. We have chosen to select a sample of what we will call ‘early-types’ on the basis of hard conservative cuts on two parameters which are available for each galaxy: the value of  $n$  obtained by fitting a single Sersic profile to the image, and the BAC probability  $p(E)$  that the object is an elliptical. We require  $n > 3$  and  $p(E) > 0.85$ . These cuts by no means select all early-type galaxies; they are simply designed to select a population which is very unlikely to be contaminated by later-types. Since our goal is to select objects of a single type, we are willing to sacrifice completeness for purity.

In support of this assertion, Table 1 shows the mix of F07 morphological types in samples which are defined by hard cuts in the BAC  $p(\text{type})$ . Table 2 reports a similar analysis which is based on the eye-ball classifications of N10 instead of F07. (Whereas BAC classifies galaxies into 4 (E,S0,Sab,Scd) morphological types, N10 use the T-Type classification ( $-5 < T < 7$ ) from the modified RC3 classifiers, and F07 use  $0 < T < 7$  in steps of  $0.5$ .) These Tables show that 91% and 86% of the resulting sample are indeed either Ell and S0. In contrast, requiring only  $n > 2.5$  (as done in the past) yields a sample in which the Ell+S0 fraction is just 62% and 56% respectively. (A small fraction of the objects

Selection	Ell	S0	Sa	Sb	Scd
EARLY-TYPES					
Selected					
P(E+S0) > 0.85 AND $n > 3$	0.57	0.29	0.14	0	0
$n > 2.5$	0.32	0.23	0.29	0.12	0.03
Missed					
P(E+S0) < 0.85 OR $n < 3$	0.14	0.43			
$n < 2.5$	0.02	0.07			
LATE-TYPES					
Selected					
P(E+S0) < 0.15 AND $n < 3$	0	0	0.15	0.41	0.39
$n < 2.5$	0.01	0.03	0.17	0.35	0.36
Missed					
P(E+S0) > 0.15 OR $n > 3$			0.79	0.35	0.17
$n > 2.5$			0.71	0.33	0.10

**Table 2.** Eyeball morphological classifications from Nair et al. (2010) who used T-Type classification using the modified RC3 classifiers. We set Ell ( $T = -5$  and  $T = -4$ ), S0 ( $T = -3$ ,  $T = -2$  and  $T = -1$ ), Sa ( $T = 0$ ,  $T = 1$  and  $T = 2$ ), Sb ( $T = 3$  and  $T = 4$ ), and Scd ( $T = 5$ ,  $T = 6$  and  $T = 7$ ).

are Irregulars, which is why the numbers do not always add up to 100%.) Clearly our selection is much purer. As a measure of its incompleteness, we also indicate the fraction of objects classified as Ells and S0s which do not make the cut. These fractions are 10% and 43% for the F07 classifications, and 14% and 43% for N10. This ‘missed’ fraction is much smaller if we only require  $n < 2.5$ , but we believe the price to pay in purity is unacceptable.

The bottom halves of the two tables show a similar analysis of BAC and  $n$  cuts which are designed to produce a pure sample of later types. In this case requiring  $n < 3$  and  $p(E) < 0.15$  yields a sample in which Sa + Sb + Scd account for 95% of the objects. If we only require  $n < 2.5$ , then Sa + Sb + Scd account for 89% of the objects so, for later types, the use of the BAC analysis does not make such a dramatic difference.

### 3.2 Dependence of the $R - L$ relation on model fitting and morphological type

The panel on the left of Figure 2 shows the  $R - L$  relation obtained for this early-type sample (i.e.  $n > 3$  and  $p(E) > 0.85$ ) based on SDSS fits to a single deVaucouleurs profile; SDSS-based `cmodel` sizes defined by Bernardi et al. (2010), which are a crude combination of separate fits to a single exponential disk and a single deVaucouleurs profile; `PyMorph` fits to a two-component deVExp model; `PyMorph` fits to a two-component SerExp model; and `PyMorph` fits to a single Sersic profile. There are clear systematic differences between these relations, with the single Sersic and deVaucouleurs models returning the relations with the largest and smallest sizes, respectively. The various two-component based relations are in good agreement except at the highest luminosities ( $M_r < -22$ ), where the sample becomes increasingly

Sample/fit	$p_0$	$p_1$	$p_2$
Sersic (Early-types)	12.8145	1.3788	0.0377
Sersic (Late-types)	8.4847	0.9092	0.0254
Sersic ( $n > 2.5$ )	8.1624	0.9821	0.0292
Sersic ( $n < 2.5$ )	4.7207	0.5601	0.0173
Sersic (P(Ell))	7.0946	0.8650	0.0262
Sersic (P(S0/Sa))	10.9232	1.2218	0.0344
Sersic (P(Sa/Sb/Scd))	13.9656	1.4694	0.0395
Sersic (P(Scd))	12.6494	1.3128	0.0352
SersExp (Early-types)	8.6032	0.9979	0.0290
SersExp (Late-types)	7.3204	0.7929	0.0226
SersExp ( $n > 2.5$ )	6.0716	0.7770	0.0242
SersExp ( $n < 2.5$ )	4.2848	0.5151	0.01615
SersExp (P(Ell))	7.4437	0.8922	0.0266
SersExp (P(S0/Sa))	9.6010	1.0903	0.0311
SersExp (P(Sa/Sb/Scd))	9.3135	1.0182	0.0286
SersExp (P(Scd))	7.8056	0.8396	0.0237
SersExp (Early-type-Bulges)	-2.0733	0.0956	0.0098
SersExp (Late-type-Disks)	6.4982	0.6934	0.0199

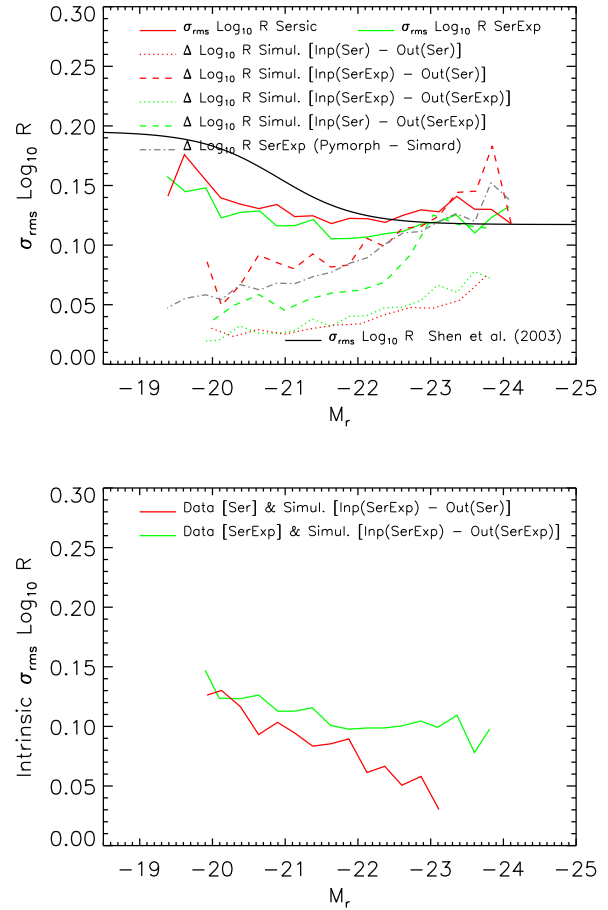
**Table 3.** Luminosity-size relation. Early-types:  $p(\text{Early-type}) > 0.85$  and  $n > 3$ . Late-types:  $p(\text{Early-type}) < 0.15$  and  $n < 3$ . Early-type-bulges: the bulge half-light radius versus the bulge luminosity for galaxies with  $p(\text{Early-type}) > 0.85$  and  $n > 3$ . Late-type-disks: the disk half-light radius versus the disk luminosity for galaxies with  $p(\text{Early-type}) < 0.15$  and  $n < 3$ .

contaminated by BCGs which are known to define steeper relations than the bulk of the population (e.g. Bernardi et al. 2007; Bernardi 2009). We argue in Appendix B that the SerExp reductions return less biased estimates of  $R$  and  $L$ .

The panel on the right shows a similar analysis of the  $R - M_*$  relation (stellar masses were computed as described in Bernardi et al. 2010 with a Chabrier IMF). Note that both  $R - L$  and  $R - M_*$  are significantly curved. (Of course, if the stellar population models used to estimate  $M_*/L$  are incorrect, or if the IMF is mass-dependent, then this will modify the curvature in  $R - M_*$ .) The curved dashed line shows a pure power law: the bottom panel shows that the deviation from this power law is substantial at  $\log_{10}(M_*/M_\odot) < 10.5$  and above  $\log_{10}(M_*/M_\odot) > 11.3$ . These are the two mass scales identified by Bernardi et al. (2011a,b).

We have repeated this analysis for a late-type sample, defined by requiring  $n < 3$  and  $p(E) < 0.15$ . Although we do not show the corresponding plots here, we again see curvature. Rather, we illustrate this in Figure 3, which compares the SerExp-based  $R - L$  relation for our way of selecting early- and late-type samples, with the more traditional cuts on  $n$  (larger or smaller than 2.5). The two ways of selecting the samples lead to very similar results, with the low luminosity early-types having smaller sizes, but defining a steeper relation, so they would cross the  $R - L$  relation of late-types at about  $M_r < -23$  (beyond which there are few late-types anyway). We have also selected an intermediate-type population by requiring  $2.5 < n < 3.5$  and  $0.2 < p(E) < 0.4$ . Notice that this sample defines the same  $R - L$  relation as when we require our early-type selection (i.e.,  $n > 3$  and  $p(E) > 0.85$ ), as well as that when we only require  $n > 2.5$ ; we return to this in Section 3.4.

We have quantified the curvature in these relations by



**Figure 4.** Top: Observed scatter around the mean  $\langle R|L \rangle$  relations for early-types based on fitting Sersic (solid red) and SerExp models (solid green) to the images. Black solid curve shows the corresponding measurement from Shen et al. (2003). Dashed and dotted curves show a number of estimates from simulations of the measurement errors (see text for details). Grey dot-dashed line in top panel shows the rms difference between PyMorph and S11 sizes (both based on fitting a two-component SerExp). Bottom: Estimate of the intrinsic scatter around the Sersic (lower, red curve) and SerExp (upper, green) derived relations for early-types, obtained by subtracting in quadrature the red-dashed and green dotted curves from the corresponding red and green solid curves shown in the top panel.

fitting to

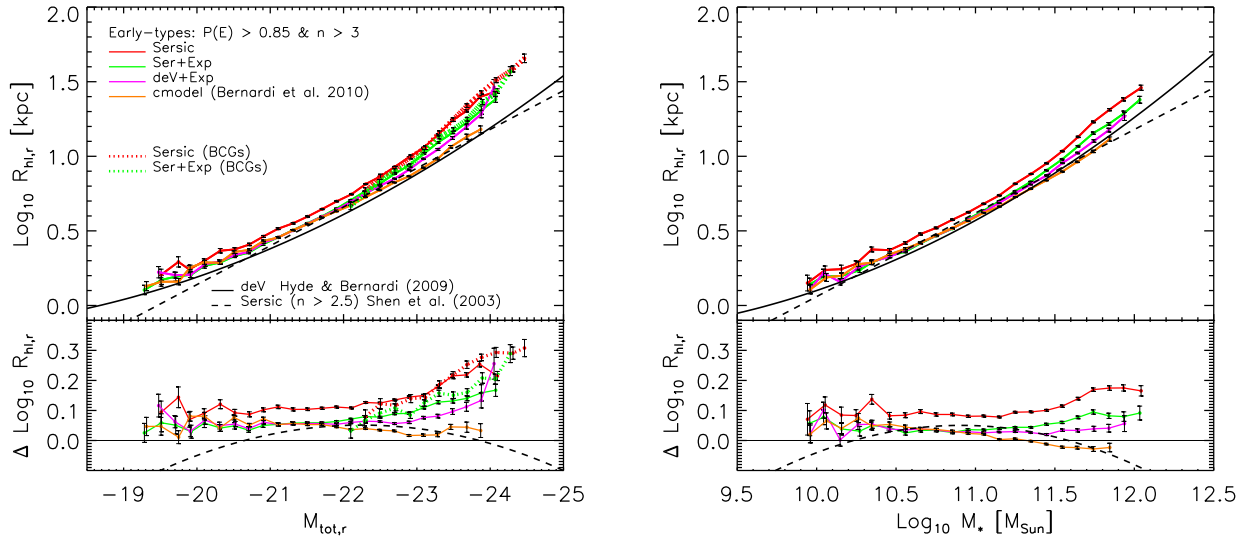
$$\left\langle \log_{10} \frac{R}{\text{kpc}} \middle| O \right\rangle = p_0 + p_1 O + p_2 O^2; \quad (1)$$

the coefficients of these fits for  $O = M_r$  and  $O = M_*$  are reported in Tables 3 and 4.

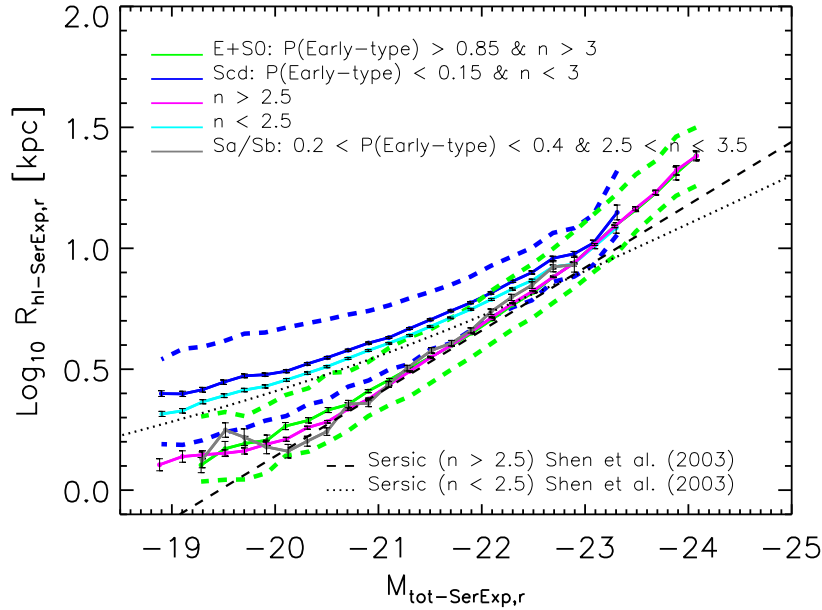
### 3.3 Scatter in $\log(\text{size})$ around the mean relation

Our analysis allows us to make two interesting statements about the intrinsic scatter around the mean  $R - L$  relation for early-type galaxies.

The top two jagged solid curves in the top panel of Figure 4 show the measured scatter around the mean  $R - L$  relation for SDSS early-types, when  $R$  and  $L$  are determined by



**Figure 2.** Dependence of derived size-luminosity (left panels) and size-stellar mass (right panels) correlations for early-type galaxies on the assumed surface brightness profile. Symbols with error bars (joined by a solid curve for clarity) show the median half-light radius in bins of absolute magnitude (left) and stellar mass (right). The SDSS fits to a single deVaucouleurs profile return a relation with the smallest sizes; our PyMorph fits to a single Sersic profile return the largest sizes. Of the relations which lie in between these two extremes, and which are almost indistinguishable at  $M < -21.5$ , the SDSS based cmodel sizes (defined by Bernardi et al. 2010) are the smallest; those based on our PyMorph fits to a two-component deVExp model are slightly larger; and those based on PyMorph fits to a SerExp model are largest. The curvature at the bright end appears to be due to an increasing incidence of BCGs, which define steeper relations (dotted lines) than the bulk of the early-type population.



**Figure 3.** Similar to previous figure, but now objects are selected using different hard cuts which define early-, late- or intermediate-type samples. Symbols with error bars (joined by a solid curve for clarity) show the median half-light radius in bins of absolute magnitude. Dashed lines show the 16th and 84th percentile. Note that this definition of intermediate's yields an  $R - L$  relation which is essentially the same as for the population with  $n > 2.5$ .

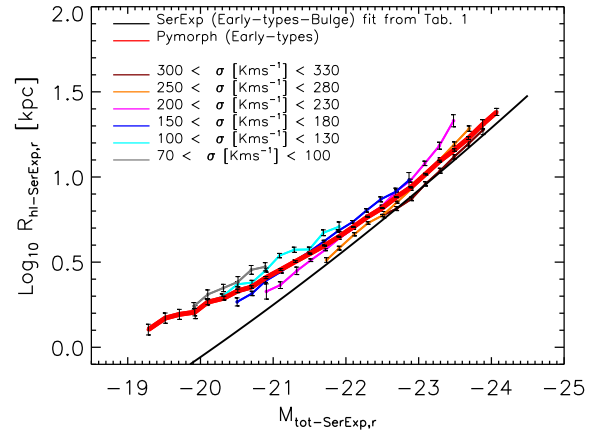
Sample/fit	$p_0$	$p_1$	$p_2$
Sersic (Early-types)	19.0933	-3.9536	0.2070
Sersic (Late-types)	13.0054	-2.6438	0.1393
Sersic ( $n > 2.5$ )	14.4995	-3.1767	0.1742
Sersic ( $n < 2.5$ )	8.6098	-1.8301	0.1014
Sersic (P(Ell))	13.6593	-2.9799	0.1635
Sersic (P(S0/Sa))	20.1092	-4.1549	0.2166
Sersic (P(Sa/Sb/Scd))	22.3082	-4.4655	0.2275
Sersic (P(Scd))	17.9815	-3.6102	0.1862
SerExp (Early-types)	13.4131	-2.9324	0.1607
SerExp (Late-types)	11.2699	-2.3026	0.1227
SerExp ( $n > 2.5$ )	12.5026	-2.7875	0.1551
SerExp ( $n < 2.5$ )	9.5210	-1.9963	0.1090
SerExp (P(Ell))	12.8394	-2.8246	0.1557
SerExp (P(S0/Sa))	19.2830	-3.9866	0.2079
SerExp (P(Sa/Sb/Scd))	18.6150	-3.7425	0.1922
SerExp (P(Scd))	11.7537	-2.3957	0.1271
SerExp (Early-type-bulges)	4.0853	-1.4159	0.0992
SerExp (Late-type-disks)	17.9763	-3.5683	0.1831

**Table 4.** Stellar mass-size relation. Early-types:  $p(\text{Early-type}) > 0.85$  and  $n > 3$ . Late-types:  $p(\text{Early-type}) < 0.15$  and  $n < 3$ . Early-type-bulges: the bulge half-light radius versus the bulge stellar mass for galaxies with  $p(\text{Early-type}) > 0.85$  and  $n > 3$ . Late-type-disks: the disk half-light radius versus the disk stellar mass for galaxies with  $p(\text{Early-type}) < 0.15$  and  $n < 3$ .

fits to a single Sersic (larger scatter) and to a SerExp model (lower scatter). This scatter is broader than the intrinsic one, because it includes a contribution from the measurement errors. For comparison, the smooth black curve shows the corresponding scatter reported by Shen et al. (2003). It is in reasonably good agreement with ours, except at the faint end, where we believe the enhanced scatter is due to increased contamination by spirals, for which the scatter is larger (as we show later).

To estimate the intrinsic scatter, we must account for the broadening due to measurement errors. We estimate the errors on the sizes from fitting to the objects in the mock catalogs used in the Appendix B, where we know the input values. (See Meert et al. 2012a for details of how the mocks were generated.) The remaining dot-dashed curve shows the rms difference between *PyMorph* and S11 sizes returned by two-component SerExp fits to SDSS images, plotted as a function of the *PyMorph* SerExp absolute magnitude for the early-type sample. This is almost certainly an overestimate of the measurement error on the sizes; we have included it just to get a sense of the overall magnitude with which systematic rather than random errors might affect the scatter in the  $R-L$  relation.

The dotted and dashed lines show our simulation-based estimates of the measurement error on the sizes for an early-type sample. The lowest dotted line shows the rms scatter in  $\log_{10} R$  around the input value if the input profile is a single Sersic, and we fit it with a Sersic. In this, and all the cases which follow, we show this scatter as a function of the fitted (as opposed to the input) absolute magnitude. The other dotted line, which lies only slightly above the previous one, shows what happens if we fit a SerExp with a SerExp. These curves certainly underestimate the full measurement error,



**Figure 5.** At fixed velocity dispersion  $\sigma$ , the  $R-L$  relation is a power law whose slope is the same for all  $\sigma$ , but whose zero-point increases as  $\sigma$  decreases.

since they are based on fits to smooth images, whereas real images may be lumpy, have spiral arms, etc.

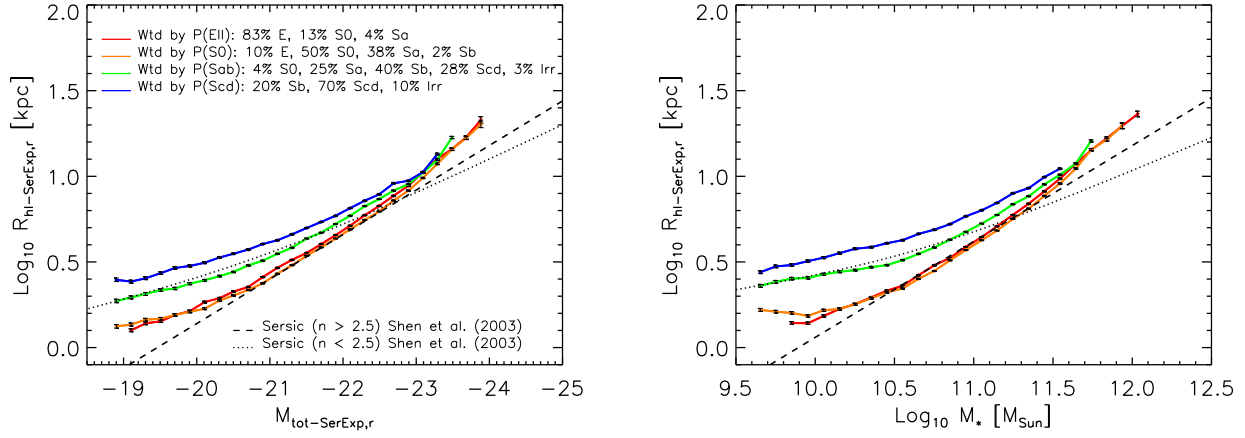
To get an idea of the magnitude of such effects, the two dashed curves show results from fitting a Sersic with a SerExp (lower) and a SerExp with a Sersic (upper). The differences between these and the dotted curves give an idea of the effect on the scatter of fitting an incorrect model to the data. The upper dashed curve is particularly interesting, in view of the fact that the SerExp model is more realistic (see Appendix B), whereas the single Sersic model is most often fit. Clearly, subtracting it in quadrature from the upper solid curve will lead to negative values at large luminosities. This is shown by the lower of the two curves in the bottom panel: at  $M_r < -23$  or so, the intrinsic scatter is consistent with zero. This, of course, does *not* mean that the  $R-L$  relation is intrinsically a line with negligible scatter. Rather, it is entirely a consequence of fitting an incorrect model.

Recently, Nair et al. (2011) have used just such an argument to claim that the  $R-L$  relation has no scatter. However, their argument is based on Petrosian sizes and luminosities; these are known to be inaccurate at large  $L$ , so the analysis above illustrates why their claim should be treated with skepticism. Indeed, the upper curve shows the result of subtracting (in quadrature) the upper dotted curve from the lower solid one, since both these are based on fitting to what we argued were more realistic models of the light profile. In this case, the intrinsic scatter is well-behaved: although it decreases steadily with  $M_r$ , it does not go negative.

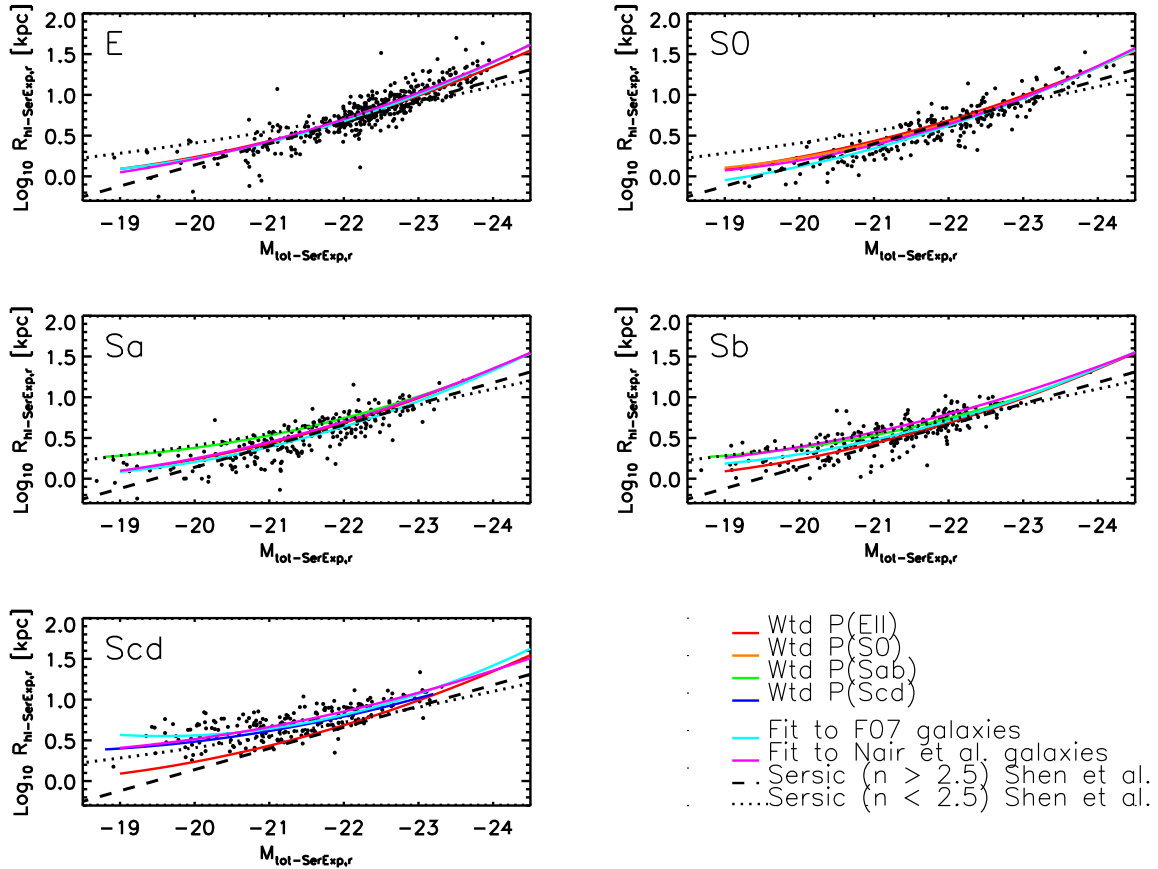
Of course, since our estimate of the measurement error is really an underestimate, it is still possible that the intrinsic scatter is smaller than we show. Therefore, we turn to what we believe is a much more effective way of showing that there is some intrinsic scatter. This method studies if the residuals from the relation correlate with other parameters, once correlations between the measurement errors have been accounted for. If they do, then there must be some intrinsic scatter.

Figure 5 shows the  $R-L$  relation for a number of narrow bins in velocity dispersion  $\sigma$ . At fixed  $\sigma$ , the  $R-L$  relation is a power law whose slope is 0.85 for all  $\sigma$  but whose zero-





**Figure 6.** Size-luminosity (left) and size- $M_*$  (right) relations obtained by weighting objects by the BAC  $p(\text{type})$ . The results from SerExp fits are shown. The low  $L$  or  $M_*$  part of the relation for Scds has the same slope as that reported by Shen et al. (2003) for their  $n < 2.5$ ; and the intermediate  $L$  or  $M_*$  part has the same slope they report for  $n > 2.5$ . Note that the relations for S0s are always indistinguishable from those for Es, and the Sab relations always lie between the E and Scd relations. Numbers in legend show the percentage of Ell, S0, Sa, Sb, Sbc and Irr galaxies classified by F07 with BAC  $P > 0.6$ . Using this selection we miss about 18% of Es, 60% of S0s, 64% of Sab (37% Sa and 27% Sb) and 56% of Scd, respectively.



**Figure 7.** Comparison of the  $R-L$  relation in the morphologically defined samples of F07 (symbols and cyan curve) and N10 (magenta), with the fits defined by the BAC of Huertas-Company et al. (2011). All relations are in good agreement for E and S0 galaxies; for comparison, the E relation is also shown in the other panels. F07 and N10 agree that Sa's define the same relation as Es and S0s, whereas Sb's are offset to larger sizes at smaller  $L$ . The HC-based results for Sab lie further from that for Es compared to those based on F07 and N10 for Sa's, but are in good agreement for Sb's; however, for Scd's they lie closer to the E relation than do F07 or N10.



point increases as  $\sigma$  decreases. In effect, this shows clearly that the scatter around the mean  $R - L$  relation correlates with  $\sigma$ ; it is not all due to measurement errors. The slope of 0.85 is consistent with previous work (Bernardi et al. 2003; Bernardi 2009); while steeper than the slope of 0.64 associated with averaging over all  $\sigma$ , it is less than unity – a fact we return to later, when we discuss the bulges of early-types (see Figure 10).

We end this subsection with the observation that the intrinsic scatter appears to be smallest for the most luminous objects. Since it is commonly believed that mergers will affect the scatter of scaling relations such as this one, our overestimate of the intrinsic scatter in the  $R - L$  relation provides a new constraint on models of how the most massive galaxies must have formed. E.g., Shen et al. (2003) argue that many minor mergers may be more consistent with the shape and scatter of the  $R - L$  relation than are few major mergers. Other work has also explored constraints which come from the scatter (Shankar et al. 2012); it will be interesting to revisit this question in light of the mass-dependence we believe we see.

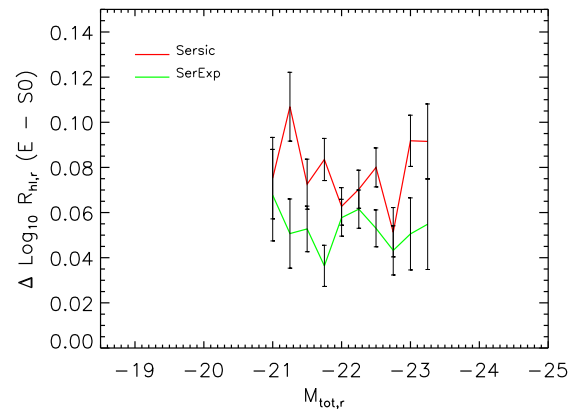
### 3.4 Dependence on morphology based on the Bayesian Automated Classifier

In the previous section we used a hard cut on the BAC probability to determine morphology. Since this is not quite in the spirit of why such probabilities were generated in the first place, Figure 6 shows the size-luminosity (left) and size- $M_*$  (right) relations obtained by weighting objects by  $p(\text{type})$  as determined by BAC. The best-fit parameters of equation (1) to these relations are reported in Table 3 and 4.

Notice that the Sab's define a relation which lies between that defined by Scd's on the one hand and Ell and S0s on the other. However, the Sab class is difficult to define (c.f. Tables 1 and 2). Figure 6 reports the percentage of Ell, S0, Sa, Sb, Sbc and Irr galaxies classified by F07 with the BAC  $p > 0.6$  – note that this is a different cut from that used in Table 1. More than a third of the objects with  $p(\text{S0}) > 0.6$  are Sa's, and about a fifth of the objects with  $p(\text{Scd}) > 0.6$  are Sb's. Conversely, of the objects which have  $p(\text{Sab}) > 0.6$ , about one third are Scd's.

To address this more closely, Figure 7 shows the  $R - L$  relations in the F07 eye-ball classified subsamples. The cyan curves show fits to these subsamples, and the magenta curves show fits based on the N10 (eyeball) classifications. The two are in quite good agreement. To emphasize the fact that the relation is different for the different subsamples, the red solid curve, which is the same in each panel, shows the BAC-based relation for  $p(\text{Ell})$ . The orange, green and blue curves (in the relevant panels) show the BAC-based relations for  $p(\text{S0})$ ,  $p(\text{Sab})$  and  $p(\text{Scd})$ . These are in good agreement with the F07 and N10 based relations for E and S0 galaxies.

Note that F07 and N10 agree that Sa's define the same relation as Es and S0s, whereas Sb's are offset to larger sizes at smaller  $L$ . This suggests that combining Sa's and Sb's into a single type may be problematic. Indeed, the BAC-based results for Sab lie further from that for Es compared to those based on F07 and N10 for Sa's, but are in good agreement for Sb's; however, for Scd's they lie closer to the E relation than do F07 or N10. These small but systematic differences



**Figure 8.** At fixed luminosity, Es tend to be about 0.06 dex larger than S0s, although this offset depends slightly on how  $R$  and  $L$  were determined.

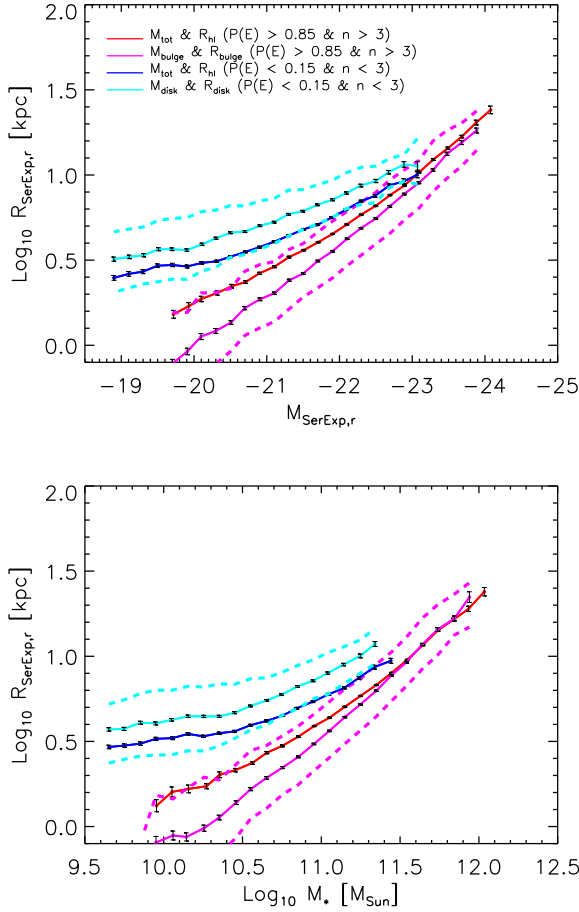
between the BAC and eye-ball based results suggest that combining Sab's into a single class results in a weighted sum of the relations defined by E's and Scd's.

Figure 7 shows that the curvature in the  $R - L$  relation is such that, for Scds, there is almost no correlation at  $M_R > -20.5$ . This flattening at low luminosities is also evident for the other morphological types, and is more pronounced in the  $R - M_*$  relation shown in the right hand panel of Figure 6 (see also Figure 9 below). Indeed, Figure 6 shows that at  $\log_{10} M_*/M_\odot < 10.5$ , even the samples weighted by  $p(\text{Ell})$  and  $p(\text{S0})$  tend to have essentially no correlation between  $R$  and  $M_*$ .

This is the same mass scale at which a number of other early-type galaxy scaling relations change qualitatively (Bernardi et al. 2011a,b). Since Figures 6 and 7 indicate that it also appears to be significant for late-type galaxies, it is interesting to ask if the other mass scale identified by Bernardi et al.,  $M_* = 2 \times 10^{11} M_\odot$ , is also significant for late-types. Figure 6 shows that, in fact, this mass scale seems to set the limit above which there are essentially no late-type galaxies. Figure 7 tells a consistent story: although there are many Es and S0s brighter than  $M_r = -23$ , there are no Sa, Sb or Scds with luminosities this large. Bernardi et al. suggested that this mass scale was associated with merger histories that were dominated by major dry mergers; since such mergers would destroy disks, the fact that we see no late-types above this mass scale is, perhaps, not surprising.

### 3.5 Small but statistically significant difference between Ellipticals and S0s

Above, we noted that there is essentially one  $R - L$  relation for E, S0 and Sa galaxies. However, our sample is large enough to detect small but significant differences within the early-type (E and S0) sample. A closer look at Figures 6 and 7 indicates that S0s are slightly smaller than Es of the same luminosity. Figure 8 shows that this offset is about 0.06 dex, although it depends slightly on how  $R$  and  $L$  were determined. This is particularly interesting in view of recent work at  $z \sim 1$ , based on the S11 reductions, which shows a similar offset of about 15% for the SDSS sample growing to



**Figure 9.** Similar to Figure 3, but now contrasting the  $R-L$  and  $R-M_*$  relations for early-types with that for their bulges, and the relation for late-types with that for their disks.

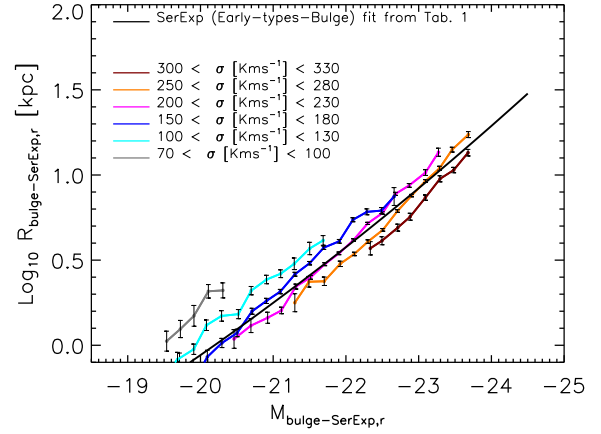
$\sim 40\%$  at  $z \sim 1$  (Huertas-Company et al. 2012). Both the sign of the trend and its evolution deserve further study.

#### 4 BULGES AND DISKS

One of the virtues of our SerExp decompositions is that it allows us to study the scaling laws of disks and bulges. Figure 9 contrasts the  $R-L$  relation for early-types with that for their bulges, and the relation for late-types with that for their disks.

The  $R_{\text{disk}} - L_{\text{disk}}$  relation runs parallel to the  $R-L$  relation for late-types;  $R_{\text{disk}}$  tends to be 0.1 dex larger than  $R_{\text{hl}}$ . That  $R_{\text{disk}} > R_{\text{hl}}$  is not surprising, since we know that late-type galaxies host small bulges which will contribute to the light at small radii. But that this should have produced a constant offset is not obvious. We address this question shortly.

The bulges are more interesting. In contrast to when the total light was used, there is almost no curvature in the relation for bulges. It is well approximated by a single power-law:  $\langle R_{\text{bulge}} | L_{\text{bulge}} \rangle \propto L_{\text{bulge}}^{0.85}$ . The amplitude of the power law is such that the relation for bulges is approximately the same as for the total at very large luminosities; as  $L$



**Figure 10.** Same as Figure 5, but showing  $R_{\text{bulge}} - L_{\text{bulge}}$  for a number of bins in total velocity dispersion  $\sigma$ . Replacing  $L_{\text{bulge}}$  with  $M_*$  of the bulge yields the same result: The relation is a power law whose slope is 1 for all  $\sigma$ , but whose zero-point increases as  $\sigma$  decreases.

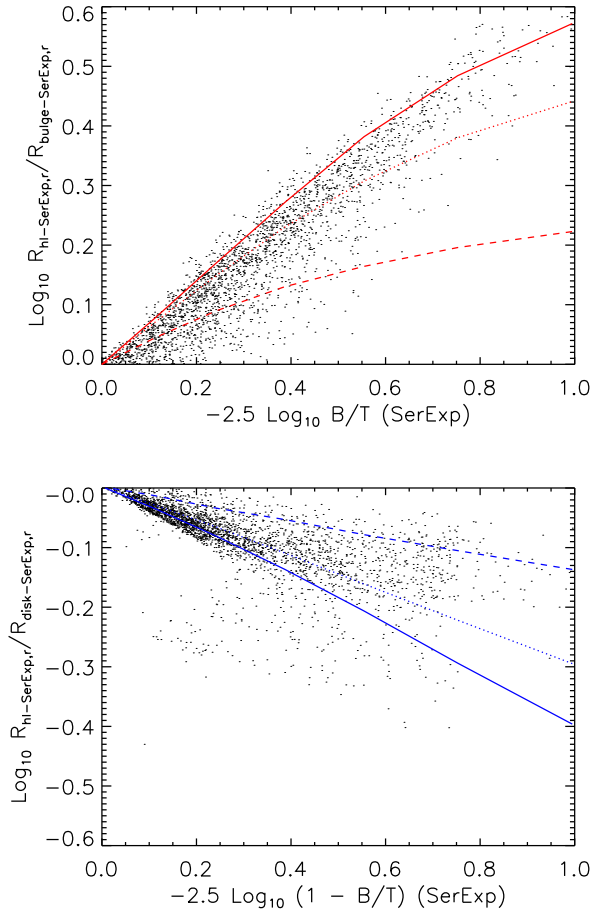
decreases, the  $R-L$  relation curves away from the  $R_{\text{bulge}} - L_{\text{bulge}}$  relation, towards larger sizes.

The power-law nature of the bulge relation suggests a picture in which the curvature in the early-type  $R-L$  relation arises as a consequence of adding a disk component to bulges. However, there is an interesting puzzle: recall that Figure 5 shows the  $R-L$  relation for a few narrow bins in velocity dispersion. This relation also has no curvature; remarkably, it runs parallel to the  $R_{\text{bulge}} - L_{\text{bulge}}$  relation, having slope  $\sim 0.85$ . To explore this further, Figure 10 shows the analogue of Figure 5: the  $R_{\text{bulge}} - L_{\text{bulge}}$  for fixed bins in  $\sigma$ . In this case, the slope is 1. Replacing  $L_{\text{bulge}}$  with  $M_{*\text{bulge}}$  makes no difference. I.e., our SerExp bulges exhibit the scaling expected from the virial theorem.

The bottom panel of Figure 9 shows the corresponding relations when plotted as a function of  $M_*$ . The  $R_{\text{disk}} - M_{*\text{disk}}$  relation is again offset from that for all late-types; both are slightly more curved than their counterparts in the panel above. At  $\log_{10} M_*/M_{\odot} < 10.5$ , the flattening of the relation with respect to the slope at large  $M_*$  is such that there is almost no correlation between  $R_{\text{disk}}$  and  $M_{*\text{disk}}$ . This flatness at the faint, low mass end is similar to that for Scds (see Figures 6 and 7 and related discussion).

The bottom panel also shows that the  $R_{\text{bulge}} - M_{*\text{bulge}}$  relation sits on top of that for early types at the largest masses, suggesting that the second component which contributes somewhat to the light contributes little to the mass. It is worth noting that this happens at the same mass scale,  $M_* = 2 \times 10^{11} M_{\odot}$ , which Bernardi et al. (2011a) noted was significant for early-types, and above which there appear to be no late-type galaxies (as is clear from this figure, as well as from Figures 6 and 7).

Of course, as we cautioned before, the conversion from  $L$  to  $M_*$  depends on  $M_*/L$ , which in turn depends on stellar population modelling as well as on an assumption about how the IMF depends on galaxy mass, so these assumptions will affect how this measured curvature – and this mass scale in particular – should be interpreted. There is another reason to be cautious: our  $M_*$  estimates assume that  $M_*/L$  for



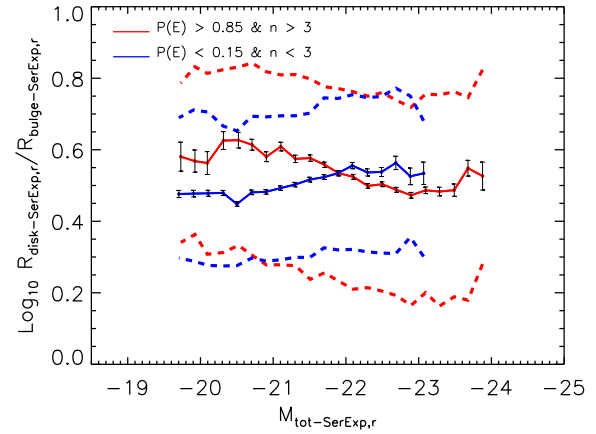
**Figure 11.** Correlation between  $R_{hl}/R_{bulge}$  and  $B/T$  for early-types (top panel) and between  $R_{hl}/R_{disk}$  and  $B/T$  for late-types (bottom panel). Although only objects with  $-21.5 > M_r > -22.5$  are shown, we see qualitatively similar behaviour at other luminosities. Dashed, dotted and solid curves show the expected scaling for  $n = 4$  bulges with exponential disks having  $R_{disk}/R_{bulge} = 2, 4$  and  $6$ .

the individual components is the same as that for the total. Since we are looking at bulges of early types and disks of late types this assumption, while crude, should not be wildly wrong (the same would not be true for, e.g., the bulges of late types). Nevertheless, one might imagine that, as a result, we slightly under(over)-estimate the mass in the bulge(disk) component.

Why are these relations for the bulge and disk components so different from those for the total light?

#### 4.1 The smallness of bulges

Suppose we start from the power-law  $R_{bulge} - L_{bulge}$  relation.  $L_{bulge}$ , with a given value of  $B/T$ , specifies a total magnitude  $M_{bulge} + 2.5 \log_{10}(B/T)$ , for which the associated half-light radius would be given by  $\log_{10} R_{bulge} + \log_{10} R_h/R_{bulge}$ . The dots in the top panel in Figure 11 show the shift in size which is associated with each shift in magnitude (itself due to  $B/T$ ), for SerExp fits to the early-type sample with  $-21.5 > M_r > -22.5$ . There is clearly a strong correla-



**Figure 12.** Dependence of the ratio of disk to bulge size on total luminosity for early- (red) and late-type (blue) galaxies.

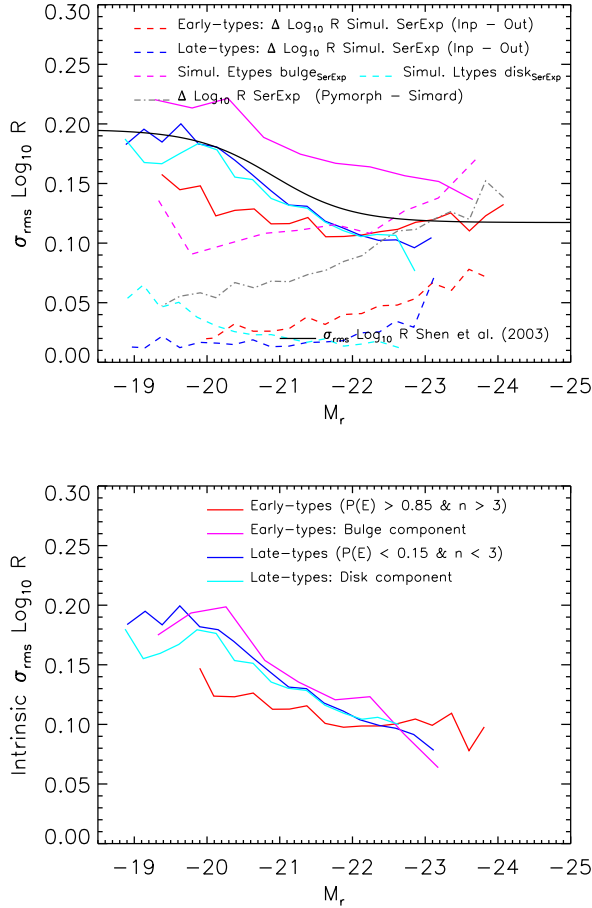
tion between  $R_{hl}/R_{bulge}$  and  $B/T$  (at this fixed  $M_r$ ) for the early-type sample. If the total is 0.55 mags brighter than the bulge ( $B/T = 0.6$ ), then the half-light radius of the total is about 0.35 dex larger than that of the bulge. What causes this?

The curves show the expected relations for a deVaucouleur bulge with exponential disk. These depend on the ratio  $R_{disk}/R_{bulge}$  (we show 2, 4 and 6) but they are independent of the total luminosity (Appendix C shows why). The dependence on  $n$  is weak (repeating the analysis with  $n = 3$  brings the curves into better agreement with the measurements;  $n = 6$  shifts in the opposite direction). Matching the data indicates that  $R_{disk}/R_{bulge} \sim 5$  at  $B/T < 0.7$ , suggesting that the correlation is caused by the fact that PyMorph uses disks with rather large scale lengths to account for the fact that a Sersic bulge is not, by itself, always a good match.

Whether or not these large scale lengths are physically reasonable is an open question, but we show in Appendix C that these tend to be objects for which the single Sersic fit returns large values of  $n > 5$ ; these extended second components do appear to be necessary to provide a good fit. Indeed, fitting SerExp images with a single Sersic profile requires large values of  $n$  if  $0.4 < B/T < 0.7$  (bottom left panel of Figure B1).

The bottom panel shows a similar analysis of the late-type sample:  $R_{hl}/R_{disk}$  as a function of  $(1-B/T)$ . Most of the sample has  $B/T < 0.2$  for which  $\log_{10}(R_{hl}/R_{disk})$  differs from zero by  $-0.05$  dex or less. Although this is in the opposite direction to the shifts for early-types (as it should be), the resulting estimate of  $R_{disk}/R_{bulge} \sim 5$  is similar. Of course, in this case, we expect  $R_{disk} \gg R_{bulge}$ , so the value of 5 does not require further explanation.

We can, of course, directly measure the ratio  $R_{disk}/R_{bulge}$  for the objects in our early- and late-type samples. Figure 12 shows that this ratio is indeed large, with only a weak dependence on  $L$ , and a somewhat larger scatter for early-types. The actual median value,  $\sim 3 - 4$ , is slightly smaller than the value of 5 we derived from the previous figure on the basis of the idealization that all galaxies were deVaucouleur bulges with exponential disks. Hence, we conclude that the differences between the relations shown in



**Figure 13.** Top: Observed (top) and intrinsic scatter (bottom) around various  $R-L$  relations as labelled (format similar to Figure 4). In all cases, our upper limit to the intrinsic scatter decreases at large luminosities; this is particularly dramatic for later-type galaxies.

Figure 9 can be traced to the fact that bulges are substantially smaller than disks.

## 4.2 Scatter

Before ending this section, Figure 13 shows our estimate of the measured and intrinsic scatter around the mean  $R-L$  relations defined by bulges and disks, and compares them with corresponding estimates for early-types and late-types. Notice that the measured scatter is substantially smaller around the early-type relation than around any of the others. Since we argued earlier that the Shen et al. (2003) early-type sample is contaminated by later-types, we believe this explains the difference between their results and ours in Figure 4. Note also that the scatter around the relation for bulges is substantially larger than for the others.

Our estimates of the intrinsic scatter (shown in the bottom panel) come from subtracting, in quadrature, the measurement errors seen in simulations (dashed lines) from the total scatter measured in the data (corresponding solid lines), following the method described in Section 3.3. For this reason, we are almost certainly overestimating the intrinsic

scatter. Nevertheless, it is interesting that for late-types, disks and bulges, our estimates indicate that the intrinsic scatter decreases at large luminosities. For early-types this decrease is less dramatic, with the scatter perhaps even levelling out at large luminosities. We believe these differences, along with the power-law nature of the bulge  $R-L$  relation, will prove to be useful for improving our understanding of how massive galaxies assembled their mass (e.g. Shankar & Bernardi 2009; Shankar et al. 2010).

## 5 SUMMARY

We used our automated image decomposition algorithm **PyMorph** to study the effects of systematics in the size-luminosity relation of galaxies in the SDSS main sample (i.e. at  $z \sim 0.1$ ) which arise from fitting different models to the images.

In Appendix A we argued that **PyMorph** returns more physically reasonable results than does the algorithm of S11 (e.g. Figures A2 and A3 and related discussion). And in Appendix B we showed that SDSS photometric information alone indicates that the majority of galaxies are not single-component systems, but have (at least) two-components. These are better modeled as a Sersic bulge plus exponential disk, rather than the traditional deVaucouleurs bulge plus exponential disk.

For objects brighter than  $L_*$ , the commonly adopted procedure – of fitting a single Sersic profile to what is really a two-component SerExp system – leads to biases. The half-light radius is increasingly overestimated as  $n$  of the fitted single component increases; it is also overestimated around  $B/T_{SerExp} \sim 0.6$ . For such objects, the assumption of a single Sersic component is particularly bad. However, the net effect on the size-luminosity relation is small, except for the most luminous tail (Figure 2).

On the other hand, fitting a realistic model is necessary to obtain sensible estimates of the intrinsic scatter around the mean  $R-L$  relation. Having done this, we showed that the scatter in sizes correlates with velocity dispersion, and the rms scatter decreases at large luminosity (Figure 4), although for early-types it may level off to a constant value of about 0.1 dex at large luminosities. This should provide tight constraints on the nature and number of mergers required to assemble the most massive galaxies.

Our Figure 6 shows one of the first use of Bayesian classifier-based weights in the estimation of the  $R-L$  scaling relation for different morphologies (e.g. Aguerri et al. 2012). We found that, even if we allow for finer bins in morphology, there seem to be only two fundamental  $R-L$  relations, both of which are slightly but statistically significantly curved (Figures 3, 6 and 7 and Tables 3 and 4). Of course, a closer inspection does reveal subtle dependences on morphology. Amongst early-types, S0s tend to be about 0.06 dex smaller than Es of the same luminosity (Figure 8). This difference smaller than the  $\sim 40\%$  reported by Huertas-Company et al. (2012) at  $z \sim 1$ . This subtle difference between Es and S0s is particularly interesting in view of the fact that the two types show very different trends as a function of age (Bernardi et al. 2010), so we expect that it, and its evolution, should yield interesting new constraints on models of how early-type galaxies assembled their stellar mass. Sim-

ilarly, amongst late-types, faint Sbs tend to be  $\sim 0.1$  dex smaller than Scds of the same luminosity, but these differences decrease as luminosity increases.

Our two-component fits allowed us to study the  $R - L$  relations for the bulge and disk components themselves. Although the  $R - L$  relations for the total light in early- and late-types are curved, the relation defined by the bulges in early-types is remarkably straight:  $\langle R_{\text{bulge}} | L_{\text{bulge}} \rangle \propto L_{\text{bulge}}^{0.85}$  (Figure 9). The relation for disks runs parallel to the  $R - L$  relation for late type galaxies, being offset upwards by about 0.1 dex. For disks, this curvature is so pronounced that, at the faint, low mass end, there is almost no correlation between  $R$  and  $L$  or  $M_*$  (Figures 7 and 9). We argued that, both for early and late type galaxies, these differences arise because PyMorph uses disk-components for which the half light radius is  $\sim 3 - 4$  times larger than that of the bulge (Figures 11, 12 and Appendix C). It is not clear if for early-types this is physically reasonable – but extended second components are clearly necessary for the SerExp fits (Figures C2 and C3).

The two mass scales,  $M_* = 3 \times 10^{10} M_\odot$  and  $M_* = 2 \times 10^{11} M_\odot$ , previously identified by Bernardi et al. (2011a,b), also appear in Figure 9. For early-types, the former, is, among other things, the mass scale at which galaxies are maximally dense. Below this scale the  $R - M_*$  relation curves upwards with respect to the power law which best describes the full range of  $M_*$  (Figures 2 and 6). Bernardi et al. suggest that this is because the disk component becomes more significant at these low masses.

The larger mass scale is where the  $R - L$  relation of early-types curves upwards with respect to the power law which best describes the full range of  $M_*$ . Bernardi et al. attribute this to a change in the assembly histories – to ones in which major dry mergers become important. So it is interesting that we find that it is at this mass scale that the bulge and total  $R - M_*$  relations become the same, despite being very different at smaller masses (Figure 9). This is particularly remarkable in light of recent work showing that early-types below this mass scale tend to be fast rotators (Cappellari et al. 2012). It may be that our SerExp bulge-disk decompositions of the images are reflecting this change in the kinematics.

Our analysis indicates that these same two mass scales are also significant for late-type galaxies. At  $M_* < 3 \times 10^{10} M_\odot$ , the  $R - M_*$  relation for late-types (and their disks) flattens significantly (Figures 6, 7 and 9); and  $M_* = 2 \times 10^{11} M_\odot$  marks the mass scale above which there are almost no late-types (Figures 7 and 9).

Given the large differences between the relation for bulges and that for early-types at smaller masses and luminosities (Figure 9), it is remarkable that the slope of the  $R - L$  relation for bulges is essentially the same as that for early-types within a fixed bin in velocity dispersion (Figure 5). Even more remarkable is the fact that, at fixed  $\sigma$ ,  $L_{\text{bulge}} \propto R_{\text{bulge}}$  (Figure 10): i.e., our SerExp bulges exhibit the scaling expected from the virial theorem. Why this should be so is an open question, but it does suggest that departures from the virial theorem scalings for the total light are entirely due to the presence of a disk component.

Finally, we find that the scatter around the mean  $R - L$  relation decreases as  $L$  increases (and similarly for  $R - M_*$ ), except for early-types, where it may flatten at 0.1 dex

(Figure 13). We expect this to provide a useful probe of how massive galaxies assembled their mass (e.g. Shankar & Bernardi 2009; Shankar et al. 2010).

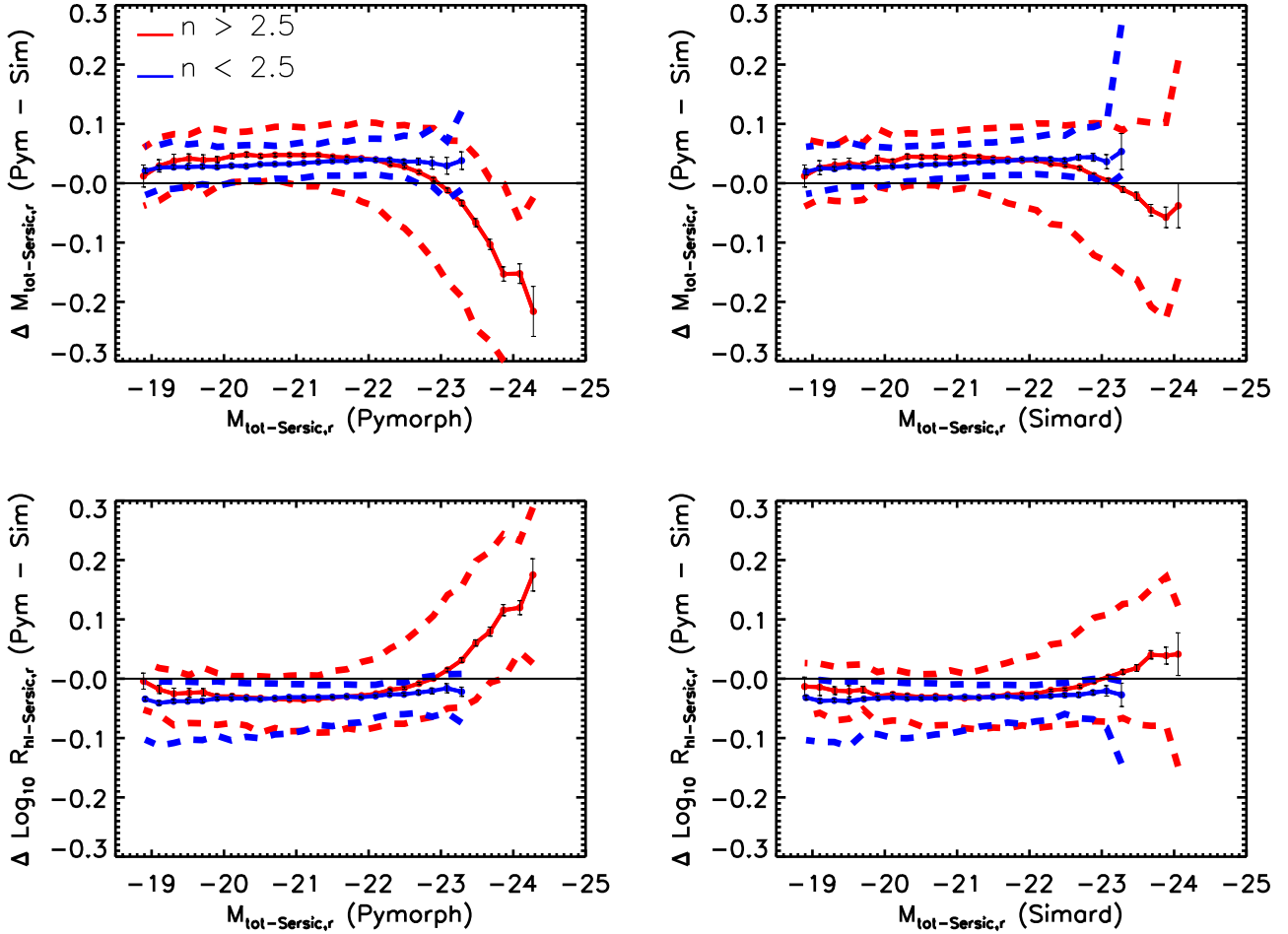
## ACKNOWLEDGMENTS

This work was supported in part by NASA grant ADP/NNX09AD02G and NSF/0908242. MB and RKS are grateful to the Meudon Observatory for its hospitality during June 2011 and 2012. FS acknowledges support from a Marie Curie grant.

## REFERENCES

- Abazajian, et al. 2009, ApJS, 182, 543
- Aguerre, J. A. L., Huertas-Company, M., Sánchez Almeida, J. & Muñoz-Tuñón, C. 2012, A&A, 540, 136
- Allen, P. D., Driver, S. P., Graham, A. W., Cameron, E., Liske, J. & de Propris, R. 2006, MNRAS, 371, 2
- Bernardi, M., et al. 2003, AJ, 125, 1849
- Bernardi M., Hyde J. B., Sheth R. K., Miller C. J., Nichol R. C. 2007, AJ, 133, 1741
- Bernardi, M. 2009, MNRAS, 395, 1491
- Bernardi M., Shankar, F., Hyde, J. B., Mei, S., Marulli, F. & Sheth, R. K. 2010, MNRAS, 404, 2087
- Bernardi, M., Roche, N., Shankar, F. & Sheth, R. K. 2011a, MNRAS, 412, L6
- Bernardi, M., Roche, N., Shankar, F. & Sheth, R. K. 2011b, MNRAS, 412, 684
- Blanton, M. R. et al. 2003, ApJ, 594, 186
- Bruce, V. A. et al. 2012, MNRAS, submitted (arXiv:1206.4322)
- Cappellari, M. et al. 2012, MNRAS, submitted (arXiv:1208.3523)
- Cimatti, A., et al. 2008, A&A, 482, 21
- deVaucouleurs, G. 1948, Annales d'Astrophysique, 11, 247
- Emsellem, E. et al. 2011, MNRAS, 414, 888
- Fukugita M., et al., 2007, AJ, 134, 579
- Hyde, J. B. & Bernardi, M. 2009, MNRAS, 394, 1978
- Huertas-Company, M., Aguerri, J. A. L., Bernardi, M., Mei, S. & Sánchez Almeida, J. 2011, A&A, 525, 157
- Huertas-Company M., Mei S., Shankar F., Delaye L., Raichoor A., Covone G., Finoguenov A., Kneib J.-P., Le Fèvre O., Povic M., 2012, MNRAS, in press (arXiv:1207.5793)
- Johnston, E. J., Aragón-Salamanca, A., Merrifield, M. R. & Bedregal, A. G. 2012, MNRAS, in press (arXiv:1202.6064)
- Meert, A., Vikram, V. & Bernardi, M. 2012a, MNRAS, submitted
- Meert, A., Vikram, V. & Bernardi, M. 2012b, MNRAS, submitted
- Nair P., Abraham R. G., 2010, ApJS, 186, 427
- Nair P., van den Bergh, S. & Abraham, R. G. 2011, ApJL, 734, 1
- Saglia, R. P. et al. 2010, A&A, 524, 6
- Sérsic, J. L. 1968, Atlas de Galaxias Australes, Observatorio Astronómico de Córdoba
- Shankar F. & Bernardi M., 2009, MNRAS, 396, L76
- Shankar F., Marulli F., Bernardi M., Dai X., Hyde J. B., Sheth R. K., 2010, MNRAS, 403, 117
- Shankar, F., Marulli, F., Bernardi, M., Mei, S., Meert, A. & Vikram, V. 2012, MNRAS, in press (arXiv:1105.6043)
- Shen S., et al., 2003, MNRAS, 343, 978
- Sheth R. K., Bernardi M., 2012, MNRAS, 422, 1825
- Simard, L., Mendel, J. T., Patton, D. R., Ellison, S. L. & McConnell, A. W. 2011, ApJS, 196, 11
- Stoughton C., et al., 2002, AJ, 123, 485
- Trujillo I., et al., 2006, MNRAS, 373, 36
- van Dokkum, P. G. et al. 2008, ApJL, 677, 5
- Vikram V., Wadadekar Y., Kembhavi A. K., Vijayagovindan G. V., 2010, MNRAS, 409, 1379





**Figure A1.** Differences between the S11 and PyMorph reductions tend to be of order 0.04 mags fainter or 0.03 dex smaller in size, except for  $M_r < -22.5$  where PyMorph tends to be bigger and brighter if  $n > 2.5$ .

## APPENDIX A: SYSTEMATIC EFFECTS IN THE SIMARD ET AL REDUCTIONS

The main text showed that the  $R-L$  relation from single-Sersic fits using PyMorph is in reasonably good agreement with that based on parameters from Simard et al. (2011). However, Figure A1 shows that, although the two algorithms return similar sizes and luminosities for objects with  $n < 2.5$  (PyMorph is about 0.03 dex smaller and 0.03 mags fainter), the PyMorph sizes and luminosities are systematically larger at large  $M_{\text{tot}}$ . This bias for the biggest galaxies is particularly evident when shown as a function of  $M_{\text{PyMorph}}$ .

Since the  $R-L$  relation of the largest galaxies is particularly timely, we would like to determine which reductions are more reliable. Figures A2 and A3 show that the S11 reductions indicate substantial recent evolution toward smaller  $n$  and  $R$  at fixed  $L$  especially at larger  $L$ . We believe this evolution is unphysical, so conclude that the S11 reductions suffer from systematic biases. No such evolution is seen in the PyMorph reductions, so we use them exclusively in the main text.

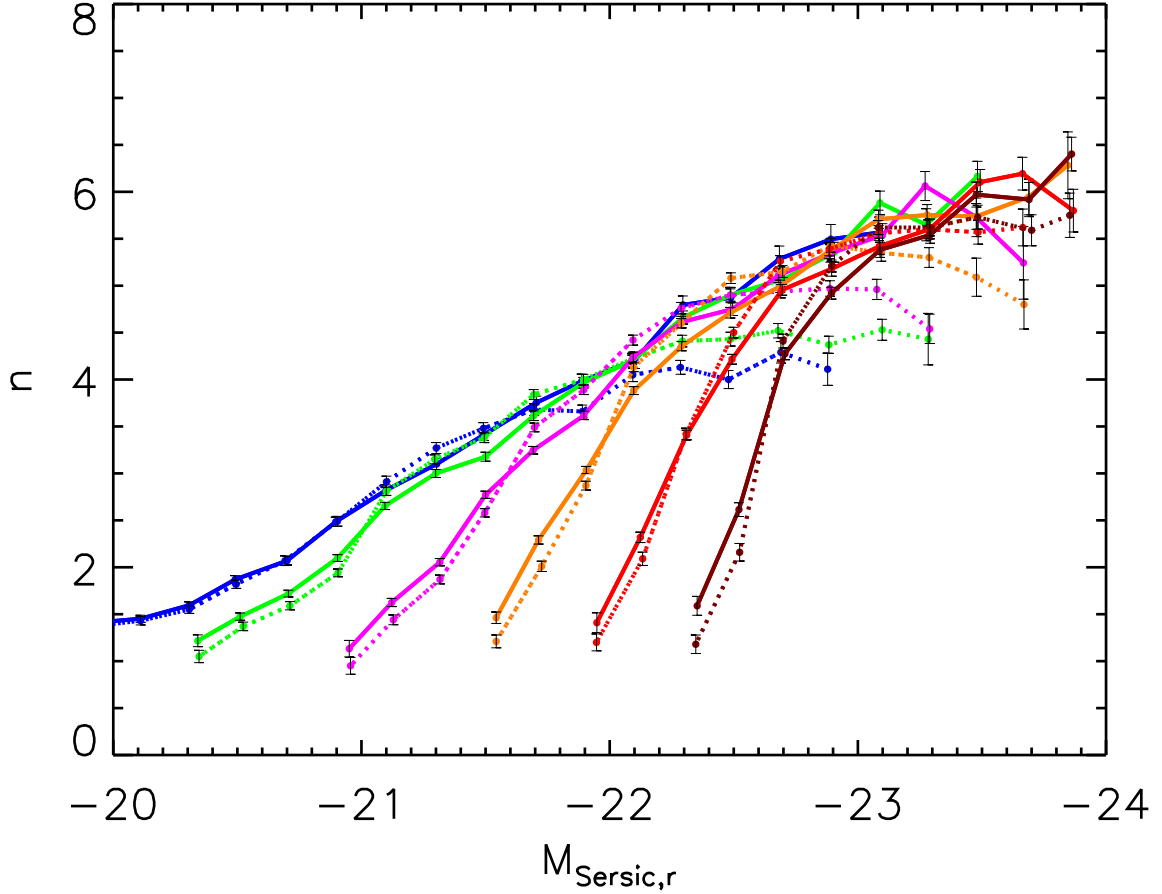
## APPENDIX B: SERSIC INDEX AND B/T RATIO IN SDSS GALAXIES: EVIDENCE FOR TWO COMPONENTS IN THE SURFACE BRIGHTNESS PROFILE

The main text considered the size-luminosity relation, noting that the derived relation depends systematically on the assumed form of the surface brightness profile. In this Appendix, we provide an analysis of the light profiles of SDSS galaxies which we believe strongly suggests that fitting to a SerExp model returns the least biased answers.

### B1 How many components?

As noted in the introduction, there has been considerable interest in developing accurate descriptions of the projected surface brightness distribution of galaxies.

One approach to this problem is to fit the free parameters of a predetermined functional form to the observed surface brightness profile. These derived free parameters (typically, these are expressed in terms of the scale which contains half the total light, and the surface brightness at this scale) are more useful if the functional form itself actu-



**Figure A2.** Our determination of the  $n - L$  relation (symbols connected by solid lines) shows little or no redshift dependence. The sudden drop in  $n$  at the faint end of each redshift sample is due to the bimodal distribution in  $n$  at each  $L$ ; it has nothing to do with evolution. Except for this, our determination shows little or no redshift dependence; in contrast, at high luminosities, the Simard et al. reductions appear to shift systematically towards smaller  $n$  as redshift decreases. We believe the implied evolution is unphysical, so conclude that the Simard et al. reductions are systematically biased.

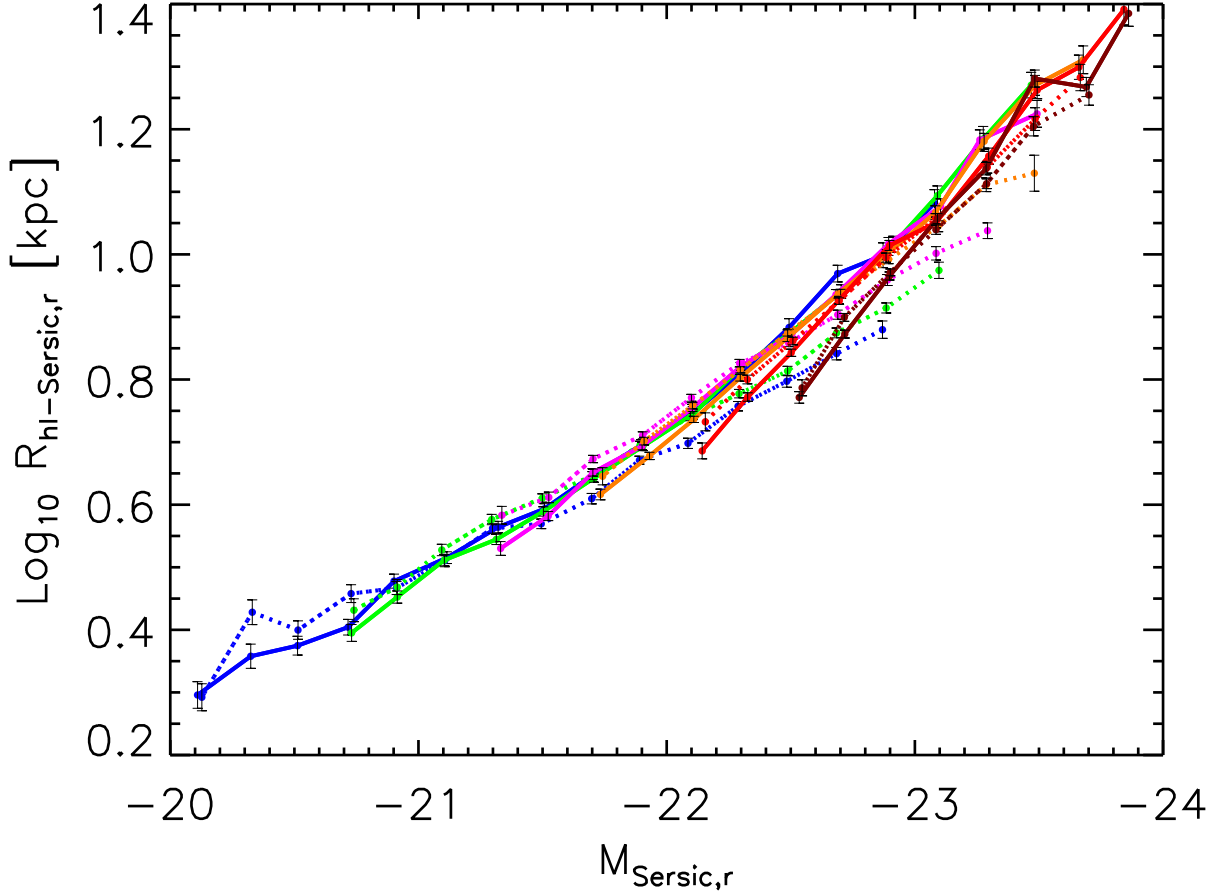
ally does provide a good fit to the profile. A simple version of this approach is to fit many different functional forms to the data, and then select the one which provides the best fit (in some suitably quantified way). For example, the Sloan Digital Sky Survey (SDSS; Stoughton et al. 2002) fits both exponential ( $I(\theta) \propto \exp(-\theta/\theta_1)$ ) and de Vaucouleurs ( $\propto \exp[-(\theta/\theta_4)^{1/4}]$ ) profiles to the image, along with an estimate of which fits better.

The exponential and de Vaucouleurs (1948) profiles are special cases ( $n = 1$  and  $4$ ) of the Sérsic (1968) profile ( $\propto \exp[-(\theta/\theta_n)^{1/n}]$ ). With sufficiently good data, it is possible to simply fit a Sérsic profile to the data, leaving the fitting procedure to determine  $n$  as well. If galaxies really are intrinsically single Sérsics with a wide range of  $n$ , then the parameters (e.g. half-light radius) returned by forcing  $n = 1$  or  $4$  in the single component fits will generally be biased. Across the population as a whole, the derived value of  $n$  spans a wide range, sometimes being as large as  $\sim 8$  or  $10$  (e.g. Simard et al. 2011 and references therein), suggesting that forcing  $n = 1$  or  $4$  is ill-advised.

Of course, it is not obvious that the light profile should be fit using a single component. The stellar kinematics in many galaxies indicate that the stars define more than one dynamical component. Examples include counter-rotating disks, as well as disk systems with bulges or bars in their centers (e.g. Emsellem et al. 2011). Evidence for more than one component is often seen in the chemical composition as well (e.g. Johnston et al. 2012). In such galaxies, it is interesting to see if the light profile also indicates the presence of more than one component.

This has motivated studies which model the observed profile as the sum of an exponential and a deVaucouleurs profile; what we will call the **deVExp** model. (Of course, since there are now more free parameters to be fit, better, higher resolution data are required. In this context, it is worth noting that Sérsic's initial motivation was to fit a functional form with fewer free parameters which would allow one to interpolate between two-component systems having varying fractions of an  $n = 4$  bulge and an  $n = 1$  disk.) It is common to report the result of such two-component fits in terms of





**Figure A3.** Similar to Figure A2, but now for the  $R - L$  relation: little or no redshift dependence is seen in our sample; in contrast, at high luminosities, the Simard et al. reductions (symbols connected by dashed lines) imply evolution towards smaller sizes as redshift decreases. We believe this implied evolution is unphysical, so conclude that the Simard et al. reductions suffer from systematic biases.

the fraction of the total light that is in the bulge (de Vaucouleurs) component: B/T. Correlations of these B/T values with other parameters (e.g. luminosity) are then used to constrain formation history scenarios.

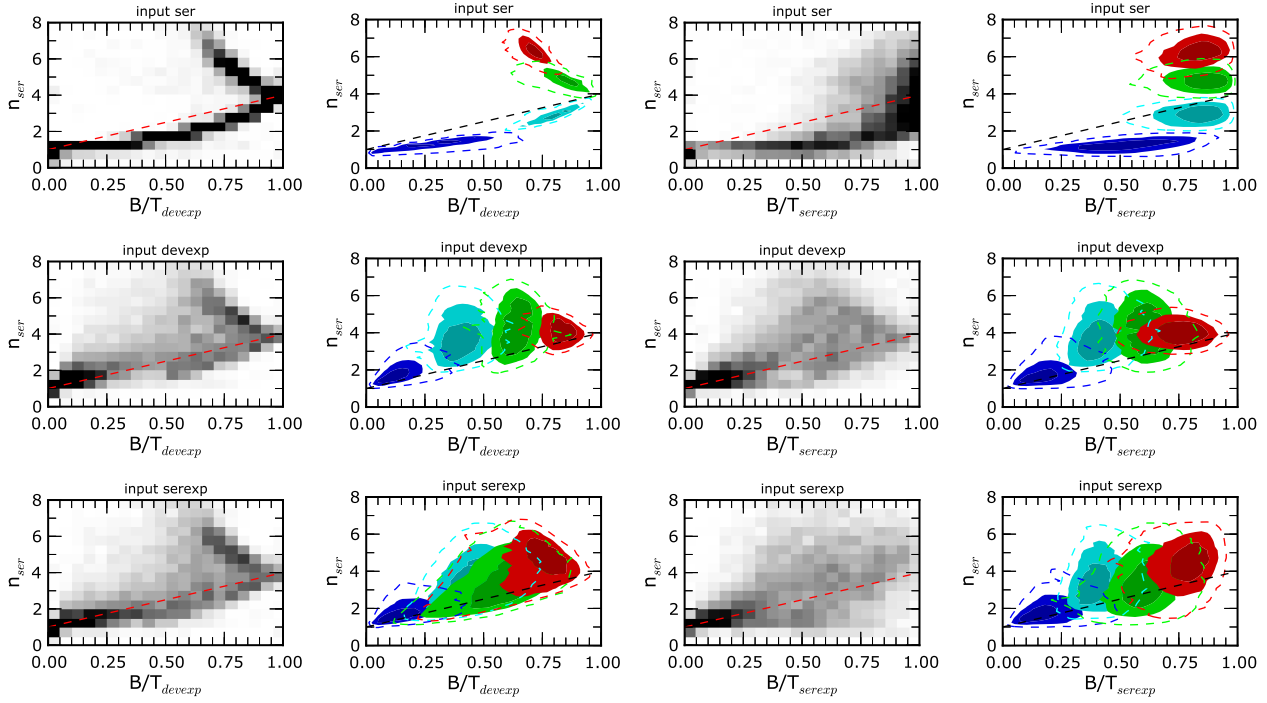
On the other hand, if galaxies really are single component Sersics, and one attempts to fit them with two component **deVExp** profiles, then one will infer an entirely spurious B/T value (the profile was, after all, just a single component). This spurious B/T will correlate with other parameters if  $n$  itself does, complicating the interpretation of such correlations. Indeed, some have argued that the evidence for two-components in the light profile is sometimes just a consequence of trying to fit what is really a single component Sersic with a linear combination of exponential and deVaucouleurs profiles. (This leaves unanswered the question of why dynamically or chemically distinct components do not leave a signature in the light.)

One way to address this question is to fit the image with the sum of two Sersic profiles, each with its own value of  $n$ , and then see if allowing for the second component does indeed provide a statistically significant improvement in the accuracy of the fit (once one accounts for the increase in

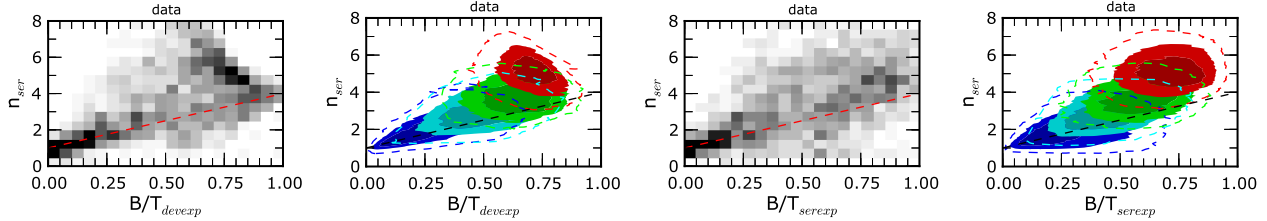
the number of fitted parameters). In what follows we will perform a slightly simpler version of this: we force one of the components to have  $n = 1$ , while leaving the other to be determined by the fitting procedure. We then provide a novel argument which indicates that this **SerExp** model is indeed a better approximation to the surface brightness profiles of real galaxies than is either a single Sersic, or the **deVExp** model.

We are not the first to have come to this conclusion; e.g. Allen et al. (2006) argued that at least half of the  $\sim 10^4$  galaxies at  $z \sim 0.1$  in the Millenium Galaxy Catalog are two component **SerExp** systems, and Simard et al. (2011) have recently performed a similar analysis of  $\sim 10^6$  SDSS galaxies. But our argument for why we believe two components are needed is new.

To gain intuition, section B2 shows the result of fitting a variety of synthetic images (generated using either a single or two-component models) with single Sersic, **deVExp** and **SerExp** profiles. This section also presents a similar analysis of real galaxies selected from the SDSS. Section B3 discusses some biases which arise from fitting the image with a single Sersic.



**Figure B1.** Fitted  $n_{\text{ser}}$  vs fitted  $B/T$  for simulated images which were generated using a single component SerSic profiles (top), or two-component **deVExp** (middle) or **SerExp** profiles (bottom). The two left columns show  $n_{\text{ser}}$ , returned by fitting a single SerSic profile to the image, versus  $B/T$ , returned from fitting a **deVExp** profile; the two right columns show the same  $n_{\text{ser}}$ , but now  $B/T$  comes from a **SerExp** fit. For each pair of columns the left column shows the density across the full sample, while the right column shows the density for four bins, colored by input  $n_{\text{ser}}$  (top; red showing larger  $n_{\text{ser}}$ ) and input  $B/T$  (middle and bottom; red showing larger  $B/T$ ).



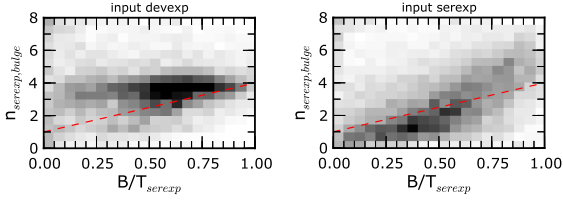
**Figure B2.** Similar to Fig. B1 but for real galaxies. Fitted (single component)  $n_{\text{ser}}$  vs fitted  $B/T$  using the two component **deVExp** fit (two left columns) and the two component **SerExp** fit (two right columns). Colors represent the probability that the galaxy is an early-type (red is highest probability).

## B2 Fits to synthetic images

In this section we show the result of using **PyMorph** (Meert et al. 2012a,b) to fit a variety of synthetic (mock) and real (SDSS) galaxies. We contrast what happens when **PyMorph** is forced to fit an image using only a single SerSic component, to when it is allowed to use two SerSic components, one with  $n = 1$  and the other free: the **SerExp** model. For the two-component fits, we first show results when  $n$  of the SerSic component is set to 4, since this corresponds to the traditional ‘deVaucouleurs bulge + exponential disk’ **deVExp** fits, and then when  $n$  is allowed to be a free parameter, determined by the fit.

In all the results which follow, the parent distribution is essentially a random subset of the SDSS DR7 main galaxy sample, which is magnitude limited to  $m_r < 17.7$ . We fit each object in this sample using three different models: a sin-

gle SerSic, a **deVExp** and a **SerExp** (see Meert et al. 2012a,b for details). We then use the best-fit parameters from these different fits to generate three synthetic images for each object. In this way, we have, in effect, three different mock SDSS catalogs (see Meert et al. 2012a for detailed tests). If galaxies were, in reality, e.g. two-component **deVExp** models, then only our **deVExp** mock catalog would be realistic – performing profile fits (e.g., using the other two models) to this catalog should return results which are similar to those when fitting to the SDSS data. Moreover, although all three catalogs will contain correlations between  $n$ , total luminosity, half-light radius, etc., these correlations are only guaranteed to be like those in the SDSS data for this (in this case, **deVExp**) mock catalog.



**Figure B3.** Parameters  $n_{\text{ser}}$  of the bulge and  $B/T$  obtained from fitting the two-component **SerExp** model to mock galaxies generated using input **deVExp** model (left) and **SerExp** model (right). In the panel on the left, the fits correctly return values of  $n_{\text{ser}} \sim 4$ ; in the panel on the right, the distribution resembles the input one: notice that this one indicates that bulges do not necessarily have  $n = 4$ .

### B2.1 Fitting to a profile which is truly a single Sersic

We begin with the case in which **PyMorph** is asked to fit what is in reality a single Sersic profile of index  $n$  (i.e. we use the mock galaxies generated using a single Sersic profile) with a single component Sersic, and with **deVExp** and **SerExp** profiles. The distribution of input  $n$  values used to simulate the mock galaxies is that which one obtains from fitting single Sersics to the parent (magnitude limited) sample. Rather than showing the fits themselves, we present our results in the parameter space of the best-fit  $n$  versus best-fit  $B/T$ . In all cases, darker shading indicates regions in the parameter space that are more heavily populated.

The top row in Figure B1 shows results for input single Sersic mock galaxies. The two panels on the left show  $B/T$  values determined from the **deVExp** fits, and the two on the right are from **SerExp** fits. We describe the **deVExp** results first. The top left panel of Figure B1 shows the distribution of the sample in best fit  $n$ – $B/T$  space, and the next panel to the right shows the result of restricting the analysis to narrow ranges of input  $n$ . The different colors show the distribution in fitted  $n$  and  $B/T$  for input  $n$  in the range  $0-2$ ,  $2-4$ ,  $4-6$  and  $6-8$  (we show the regions which enclose 25%, 50% and 75% of the points). Comparison with the values along the y-axis shows that **PyMorph** correctly returns the input  $n$  values.

The distribution in the  $n$ – $B/T$  plane is clearly non-trivial. For  $n < 4$  there is a tight correlation between the value of  $n$  returned by the single component and  $B/T$  from the deVaucouleurs-exponential fit:  $B/T \rightarrow 1$  as  $n \rightarrow 4$ . But as  $n$  increases beyond 4,  $B/T$  begins to decrease again. I.e.,  $B/T$  is not a monotonic function of  $n$ . Since the **deVExp** profile only has  $n = 1$  or  $n = 4$  components, to fit  $n > 4$  profiles **PyMorph** requires more and more of an exponential-like component, i.e.  $B/T$  decreases. (The figure does not show this, but the fit returns bulge half-light radii which are ever smaller fractions of the half-light radius of that of the input Sersic profile.) As a result, for  $1/2 < B/T < 1$ , the distribution of  $n$  at fixed  $B/T$  appears bimodal. This shows that, unless one is certain that large values of  $n$  do not occur in nature, then, especially around  $B/T \sim 0.7$ ,  $B/T$  values may be misleading, if not meaningless.

The two panels on the right show the corresponding distribution for **SerExp**; they are clearly different from those for **deVExp**. This is primarily because **PyMorph** correctly assigns the entire profile to the bulge (Sersic) component, except

when the input  $n \sim 1$ , since then which of the two  $n = 1$  components should be called the bulge is ambiguous. (We have checked that, when  $n \sim 1$  and  $B/T < 1$ , then the half-light radius of the ‘bulge’ component is indeed the same as that of the total: i.e., the two components differ only by the value of  $B/T$ .) The fact that  $B/T$  is not exactly equal to unity is a measure of the error in  $B/T$  which comes from the extra degree of freedom associated with having a second component with which to fit the profile.

### B2.2 Fitting to a profile which is truly a deVExp

The second row shows results when the input profile used to simulate the mock galaxies is a two component **deVExp** model (the distribution of input  $B/T$  values is obtained from fitting **deVExp** models to the SDSS parent magnitude limited sample). This two-component profile is then fit with a single Sersic to get  $n$ ;  $B/T$  comes from fitting a **deVExp** model (two panels on left) or a **SerExp** model (two panels on right). The overall (grey-scale) distributions are rather different than in the corresponding panels in the top row. This is the first hint that the distribution of fitted  $n$ – $B/T$  can be used as a diagnostic of the true profile shape. Different colors show results for narrow bins in input  $B/T$ ; these indicate that **PyMorph** indeed returns the correct values when it fits the right model. The additional freedom when fitting a **SerExp** profile to what is really a **deVExp** means that, in the panel on the far right, the distribution of fitted  $B/T$  at fixed input  $B/T$  is slightly broader than when fitting a **deVExp**.

### B2.3 Fitting to a profile which is truly a SerExp

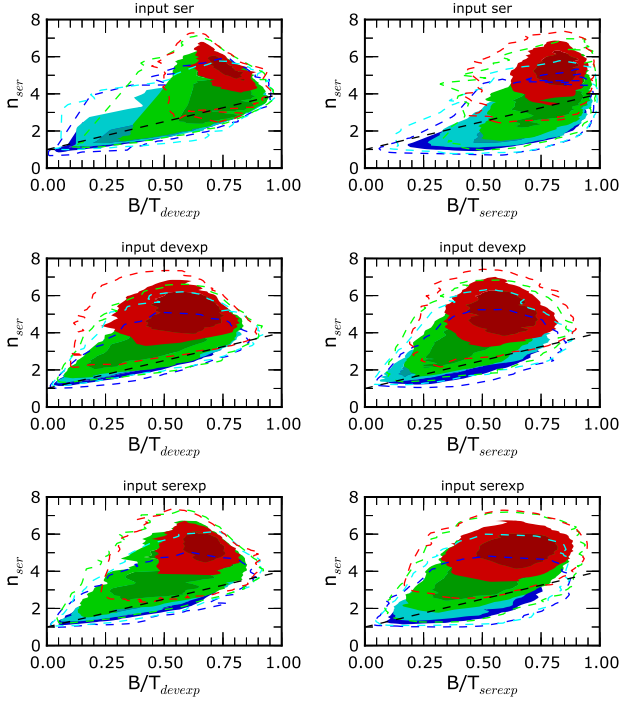
Finally, the bottom row shows results when the input model used to simulate the mock galaxies was a **SerExp** (with  $n$  and  $B/T$  values chosen from fitting the SDSS parent sample to a **SerExp** model). The results here differ from those in the row above in subtle ways, perhaps most appreciably in the upper right corner (large fitted  $n$  and  $B/T$ ) of the bottom right plots.

In this case, we also show (Figure B3) the  $n_{\text{bulge}}$ – $B/T$  plane, where both  $n_{\text{bulge}}$  and  $B/T$  come from fitting a **SerExp** model to mock images generated using input **deVExp** (left) and **SerExp** (right) profiles. The panel on the left shows that **PyMorph** correctly returns  $n \sim 4$  when it should; we have checked that the distribution in the panel on the right is similar to the input one, again suggesting that **PyMorph** is working well (Meert et al. 2012a).

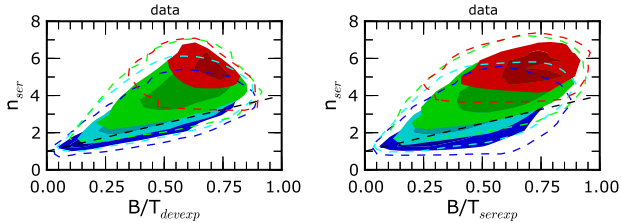
### B2.4 Fitting to SDSS images

Figure B2 shows a similar analysis of SDSS images. In the two panels on the left,  $n$  comes from fitting a single component Sersic, and  $B/T$  from fitting a two-component **deVExp**. In the panels on the right,  $B/T$  comes from fitting a two-component **SerExp**. Notice that the gray scale plots are very unlike those in the top row of Figure B1, and most like those in the bottom row. This suggests that SDSS galaxies are almost certainly not single-component systems.

In addition, of the two-component models, the **SerExp** model appears to be more like the data than is the **deVExp**. This is because, when  $B/T$  comes from fitting a **SerExp**,



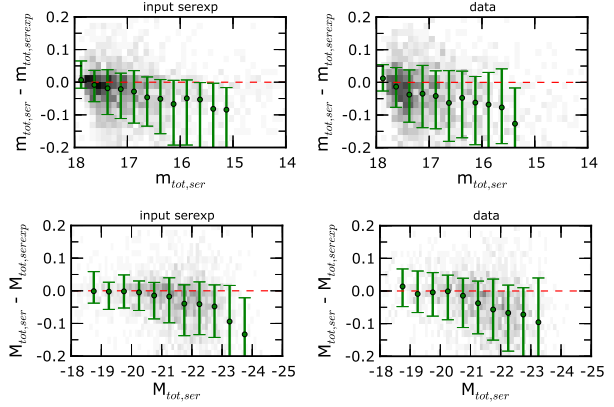
**Figure B4.** Fitted  $n_{\text{ser}}$  vs fitted  $B/T$  for simulated galaxies that are assumed to be single Sersic profiles (top), two-component deVExp profiles (middle) and two-component SerExp profiles (bottom). In all cases, the y-axis shows  $n_{\text{ser}}$  returned by fitting a single Sersic profile to the image. In the left column,  $B/T$  is obtained from fitting a two-component deVExp model; the right column,  $B/T$  is determined from fitting a SerExp model. The density is shown in four bins colored by output absolute magnitude:  $-24 < M_r < -23$  (red),  $-23 < M_r < -22$  (green),  $-22 < M_r < -21$  (cyan),  $-21 < M_r < -20$  (blue).



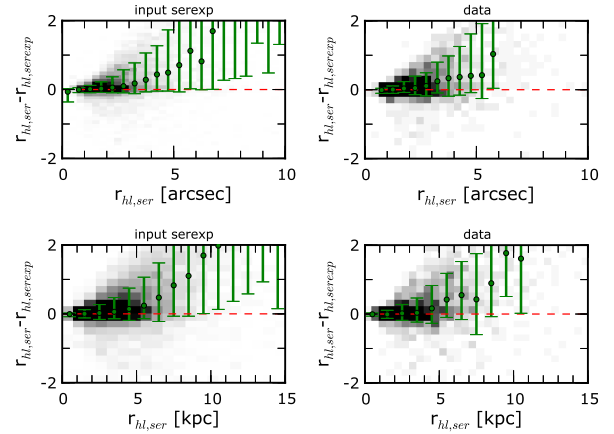
**Figure B5.** Similar to Fig. B4 but for real SDSS galaxies.

then the SDSS data (third panel from left) populate the large  $n$ – $B/T$  corner which input SerExp models also fill, but input deVExp models do not (c.f. Figure B1). There is a more subtle difference when  $B/T$  comes from the deVExp fit (left most panels) in Figures B1 and B2: the SDSS shows a rather well-defined ridge at the boundary of the large  $n$ – $B/T$  corner, which appears to be more separated from the peak at small  $n$ . This separation is more apparent for the input SerExp models than for input deVExp.

Since we cannot classify the objects by the true value of  $n$  or  $B/T$ , the colors (contours in Fig. B2) show the result of restricting the analysis to objects which are most likely to be ellipticals (red) to least likely (blue), as determined by Huertas-Company et al. (2011). This shows that the ellipti-



**Figure B6.** Comparison of total apparent magnitude (top) and luminosity (bottom) returned from single Sersic and SerExp fits to simulated SerExp (left) and real SDSS (right) galaxies. The error bars show the  $1\sigma$  rms scatter around the median.



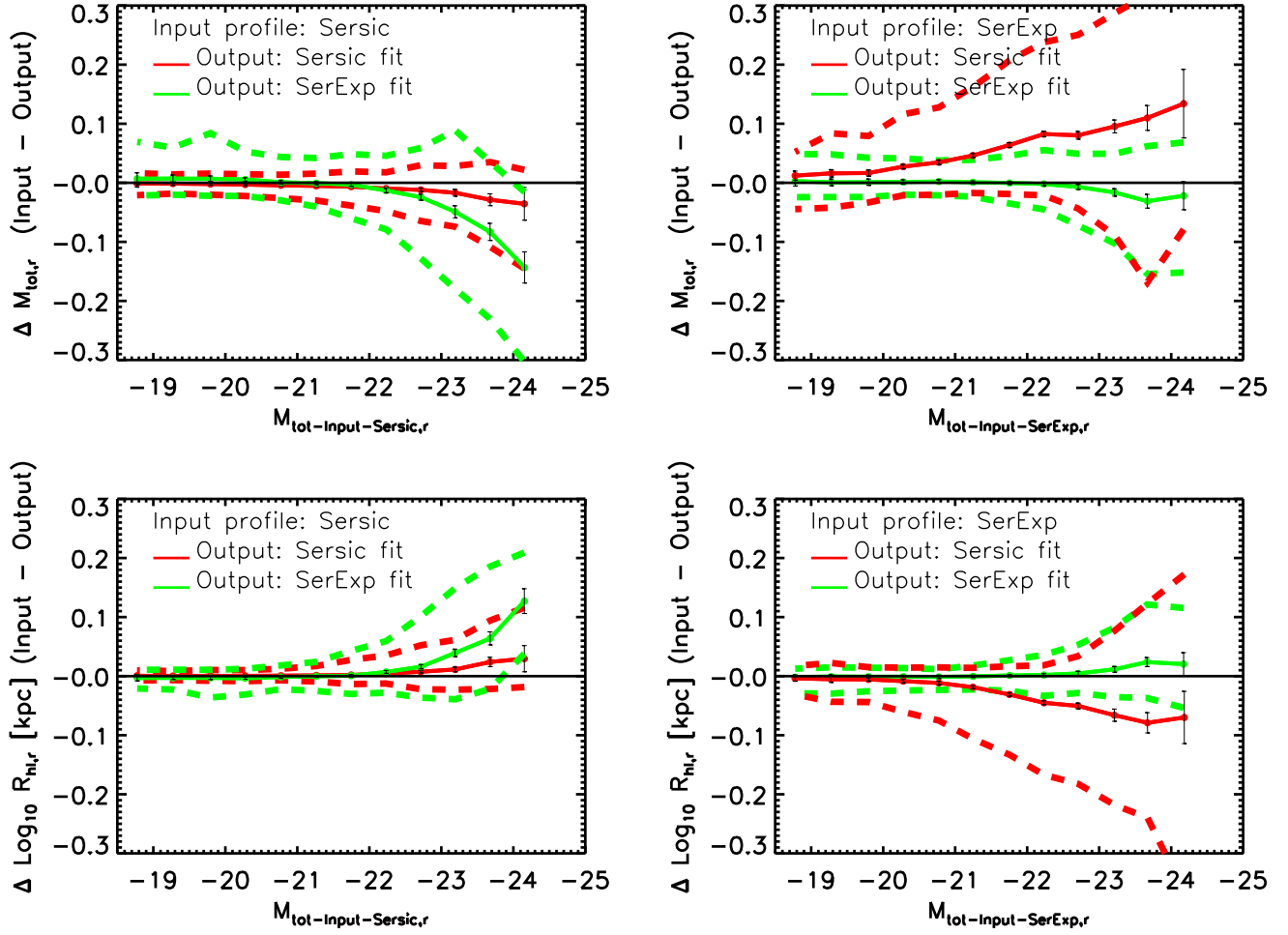
**Figure B7.** Comparison of angular (top) and physical (bottom) half-light radii returned from single Sersic and SerExp fits to simulated SerExp (left) and real SDSS (right) galaxies. The error bars show the  $1\sigma$  rms scatter around the median.

cals do indeed have large values of  $n$ , and spirals the lowest, as expected.

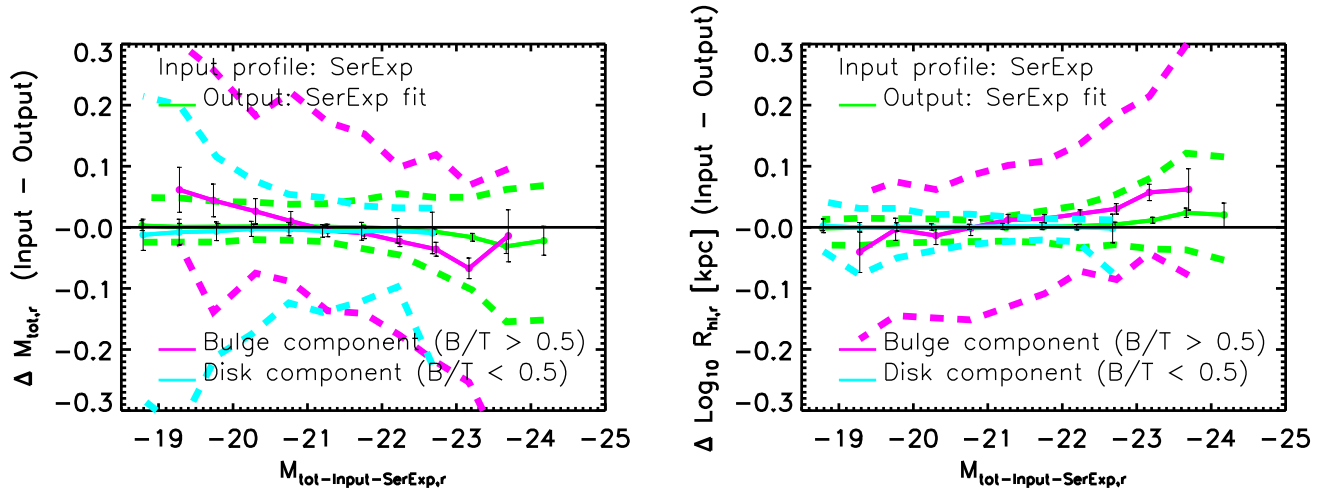
To provide a slightly more straightforward comparison between simulations and data, we have considered the  $n$ – $B/T$  distribution for objects in narrow bins in (output) luminosity. Figures B4 and B5 show results in simulations (the same fits used for Figure B1) and in the SDSS (cf. Figure B2), respectively. These too indicate that the two-component models are more like the data, with the SerExp marginally favoured (the two panels in Figure B5 look more like the bottom than the middle panels of Figure B4).

### B3 Biases from fitting single Sersic profiles

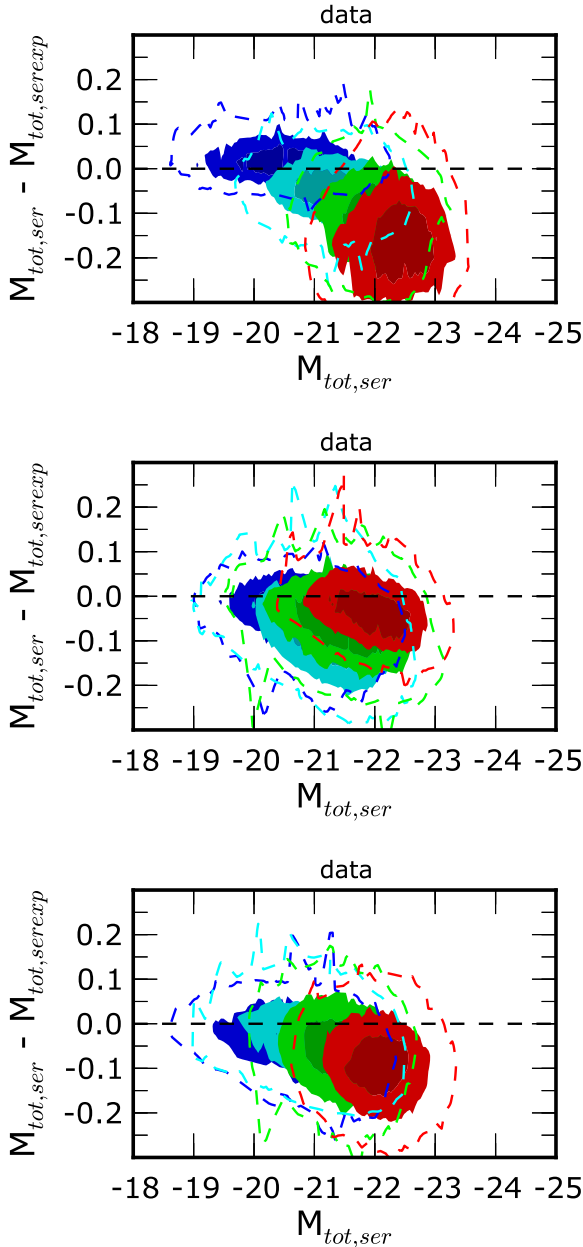
The analysis above shows that a single component Sersic profile is not as good a description of SDSS galaxies as one with two-components. Since such single component fits are much simpler to perform, and are commonly used, it is interesting to ask if they lead to significant biases in commonly used parameters. E.g., one might expect the total light to



**Figure B10.** Biases in the estimated luminosities and sizes which come from fitting single Sersic and two-component SerExp profiles to images which are really pure Sersics (left) and two-component SerExps (right).



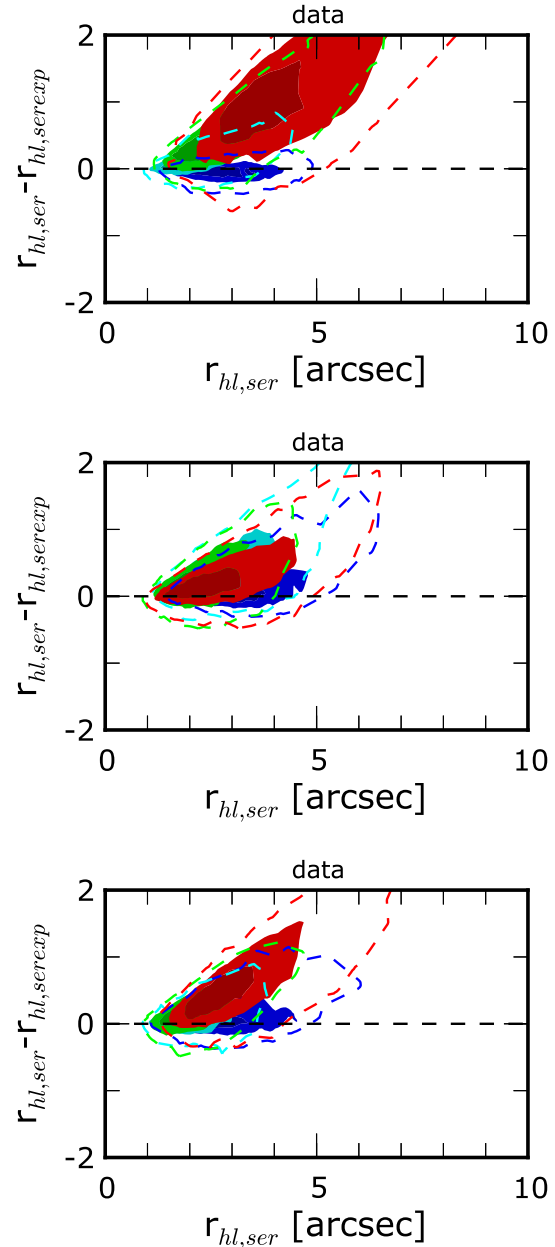
**Figure B11.** Biases in the estimated luminosities (left) and sizes (right) of the total (green), bulge (magenta) and disk (blue) components in SerExp fits to SerExp images. The estimated total and disk components are usually unbiased, whereas the bulges tend to be too big and too bright at large luminosities.



**Figure B8.** Comparison of luminosities returned from single Sérsic and **SerExp** fits to real SDSS galaxies, color coded by best-fit  $n$  (top), best-fit B/T (middle), and  $p(ell)$  (bottom).

be a reasonably robust quantity, so different models for the shape of the profile may still return consistent values of  $L_{tot}$ . The top left panel of Figure B6 shows that the apparent magnitudes returned by single Sérsic fits to the objects in our **SerExp** mock catalog are quite accurate, with a tendency for the Sérsic fits to return an overestimate by about ten percent at the bright end. The top right panel shows that a similar comparison for the objects in the SDSS produces similar results. The bottom panels show the impact of these small biases on the inferred luminosities.

Figure B7 shows a similar analysis of the half-light radii: the single Sérsic fit tends to overestimate the sizes by about ten/fifteen percent, particularly for the largest objects. The



**Figure B9.** Comparison of angular half-light radii returned from single Sérsic and **SerExp** fits to real SDSS galaxies, color coded by best-fit  $n$  (top), best-fit B/T (middle), and  $p(ell)$  (bottom).

largest and/or most luminous galaxies tend to have large  $n$  and/or intermediate to large B/T. Therefore, this bias is worst for objects that are likely to be ellipticals. We show this explicitly in Figures B8 and B9.

We have presented a novel diagnostic of whether or not the surface brightness profiles of galaxies are better thought of as having one or two components. The method works by fitting a number of single and two-component models to the image, and then studying the distribution in the  $n$ -B/T plane defined by the Sérsic index  $n$  associated with the single component fit and the ratio B/T of bulge to total light in the two component fit. The way SDSS galaxies populate this plane suggests that they are not single component Sérsic



systems. Rather, their distribution in  $n$ -B/T is more similar to that expected of two-component systems, with a Sersic + exponential model faring somewhat better than the traditional deVaucouleurs bulge with exponential disk model (Figures B1 and B2). I.e., in bulge dominated systems, allowing  $n \neq 4$  provides a significantly improved fit. Indeed, we even find bulges with  $n > 4$  in the SDSS (Figure B3).

Our conclusion that the **SerExp** model is preferred is consistent with a recent analysis of the MGC, indicating that at least half of the galaxies at  $z \sim 0.1$  are two component **SerExp** systems (Allen et al. 2006). Forcing the light profiles to be fit by a single Sersic profile leads to biases in the inferred total luminosities and half-light radii of galaxies: both are overestimated by about ten percent, especially at large luminosities and sizes. The bias is dominated by objects with large fitted  $n$  and/or intermediate values of fitted B/T; these are objects that are likely to be early-type, but have a significant exponential component, so the assumption of a single profile is particularly bad. In contrast, objects that are likely to be late-type are unbiased.

These biases have a small systematic effects on the size-luminosity correlation of objects that are likely to be early-types, and are presented in the main text.

### APPENDIX C: CORRELATION BETWEEN $R_{\text{BULGE}}/R_H$ OR $R_{\text{DISK}}/R_H$ AND B/T

Figure 11 of the main text showed a correlation between the bulge(disk) to total size and B/T. This correlation arises because the Sersic profile is

$$I(r) = I_n \exp[-(r/r_n)^{1/n}] \quad (\text{C1})$$

so the ratio of the light within  $r$  to the total light in the profile is

$$\frac{L_n(< r)}{L_n} = \frac{\int_0^{r/r_n} dx x \exp(-x^{1/n})}{\int_0^\infty dx x \exp(-x^{1/n})} = \gamma_{2n} \left[ 0, (r/r_n)^{1/n} \right] \quad (\text{C2})$$

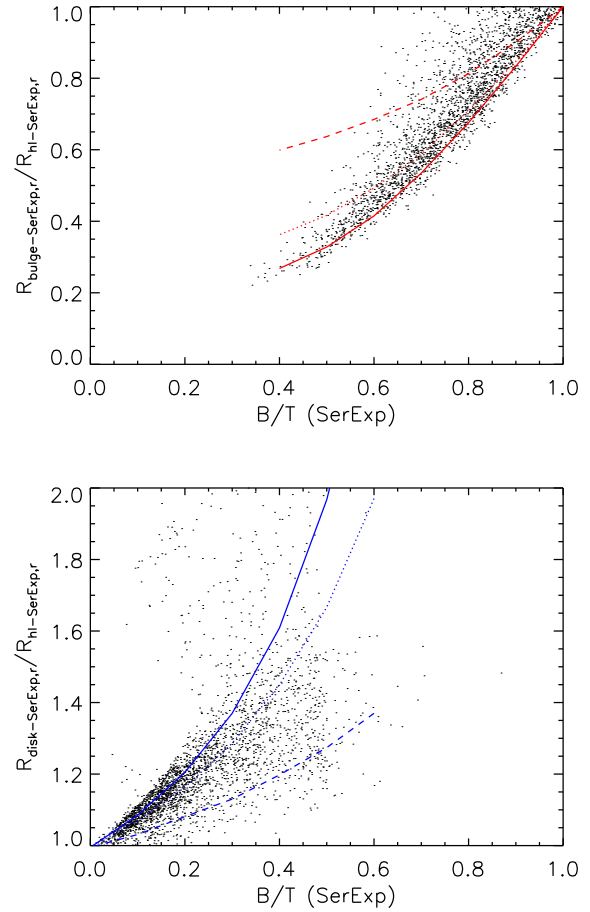
where  $\gamma_{2n}$  is the incomplete Gamma function. (For integer  $n$ , it can be written in terms of  $\exp[-(r/r_n)^{1/n}]$  times a polynomial of degree  $2n - 1$  in  $(r/r_n)^{1/n}$ .) Therefore, the half-light radius  $r_h$  of a SerExp profile satisfies

$$\frac{1}{2} = \frac{B}{T} \gamma_{2n} \left[ 0, \left( \frac{r_h}{r_n} \right)^{1/n} \right] + \left( 1 - \frac{B}{T} \right) \gamma_2 \left[ 0, \left( \frac{r_h}{r_1} \right)^{1/n} \right]. \quad (\text{C3})$$

For a given B/T, the right hand side is a function of  $r_h/r_n$  and  $r_h/r_1 = (r_h/r_n)(r_n/r_1)$ , so it defines a different curve for each  $r_n/r_1$ , where  $r_n = r_{\text{bulge}}/(1.992n - 0.327)$  and  $r_1 = r_{\text{disk}}/1.67$ . Note that the curves are independent of luminosity  $L$ ; therefore  $L$  dependence only enters if the distribution of  $r_n/r_1$  and/or B/T depend on  $L$ .

Figure C1 shows  $R_{\text{bulge}}/R_{\text{hl}}$  as a function of B/T for the early- (top) and late-type (bottom) samples defined in the main text for galaxies with  $-21.5 > M_r > -22.5$ ; results are similar at other luminosities. The curves show the predicted relations (equation C3) for a deVaucouleur bulge ( $n = 4$ ) with exponential disk. These depend on the ratio  $R_{\text{disk}}/R_{\text{bulge}}$ , for which we have chosen 2, 4 and 6.

The top panel shows a very strong correlation between  $R_{\text{bulge}}/R_{\text{hl}}$  and B/T (at this fixed  $M_r$ ) for the early-type



**Figure C1.** Correlation between  $R_{\text{bulge}}/R_{\text{hl}}$  and B/T for early-types (top panel) and between  $R_{\text{disk}}/R_{\text{hl}}$  and B/T for late-types (bottom panel). Although only objects with  $-21.5 > M_r > -22.5$  are shown, we see qualitatively similar behaviour at other luminosities. Dashed, dotted and solid curves show the expected correlation for  $n = 4$  bulges with  $R_{\text{disk}}/R_{\text{bulge}} = 2, 4$  and  $6$ .

sample. Clearly, if 20% of the light is in a disk component, then the size is affected by at least this fraction. The well-known correlation between  $L$  and B/T, and the fact that early-types span a large range of B/T, means that the bulge and early-type size-luminosity relations can be quite different. It is perhaps surprising that the half-light radius of the disk component is typically more than 3-5 times larger than that of the bulge, particularly at  $B/T < 0.7$ . We return to this shortly.

The bottom panel shows  $R_{\text{disk}}/R_{\text{hl}}$  and B/T for the late-type sample. Most of the sample has  $B/T < 0.2$ . Comparison with the smooth curves indicates that  $R_{\text{disk}}/R_{\text{bulge}} \sim 5$  for most of the sample. In this case we do expect the disks to be substantially larger than the bulges, so the results are sensible.

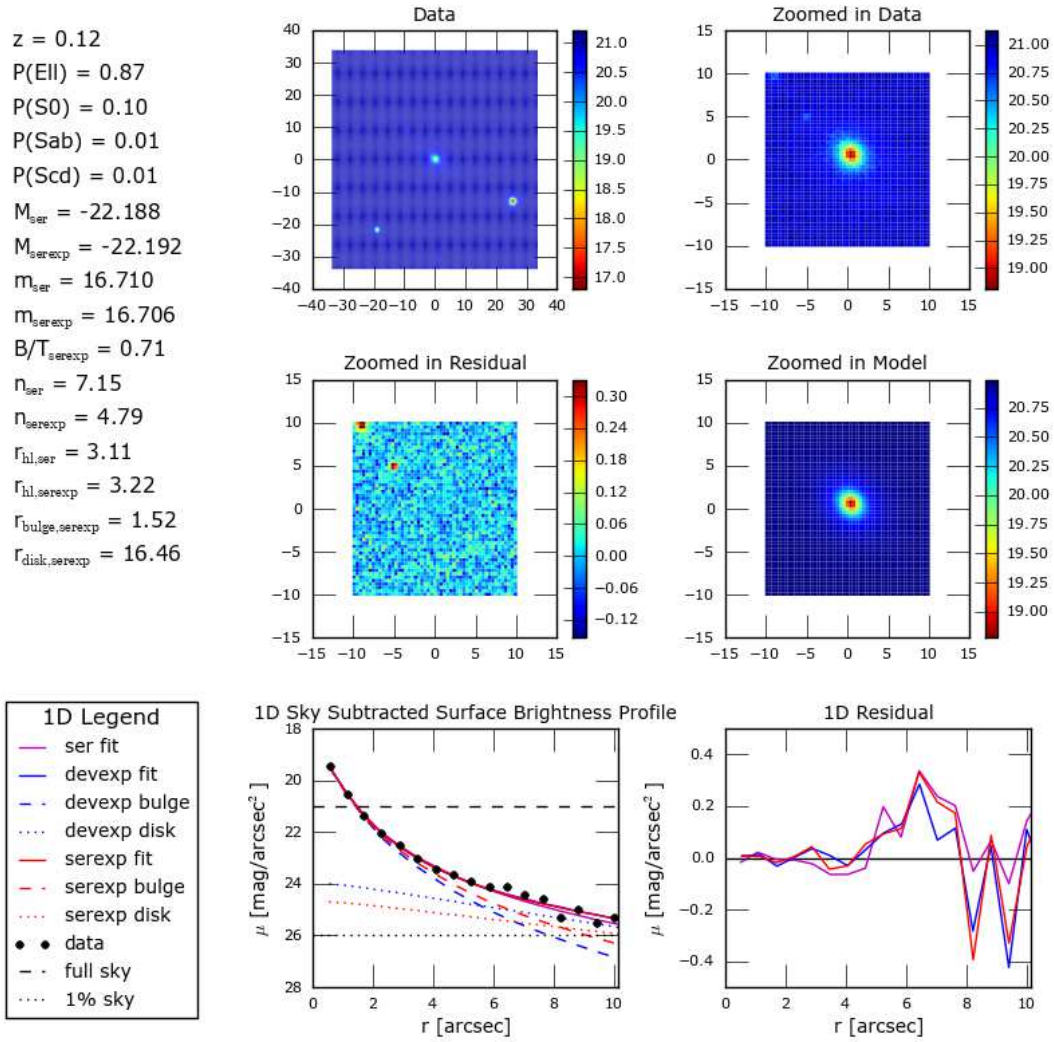
Figure 11 in the main text shows this same information in a different format, which allows for a more direct understanding of the impact this correlation has on the relations shown in Figure 9. And Figure 12 in the main text shows that  $R_{\text{disk}}/R_{\text{bulge}}$  is indeed substantially larger than unity.

To address the question of large  $R_d/R_b$  in our early-



type sample, particularly at smaller B/T, Figures C2 and C3 show two examples. The format in both cases is the same: The top left panel shows a  $\sim 20$  arcsec field centered on the object, to get an idea of whether or not the object is in a crowded field. The top right panel provides a closer look at the object. The panel just below it shows the best-fit SerExp model, and the middle left panel shows residuals from this fit. The bottom left panel shows the one-dimensional surface brightness profile, and our Sersic (solid magenta), deVExp (solid blue) and SerExp (solid red) fits; dotted and dashed curves show the corresponding disk and bulge components. Bottom right panel shows the associated residuals. The legend along the left shows the values of many quantities returned by the fits, and other information, such as the BAC  $p(\text{type})$ , for the object.

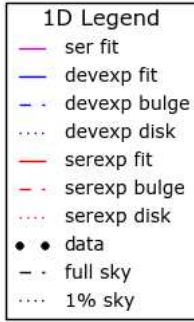
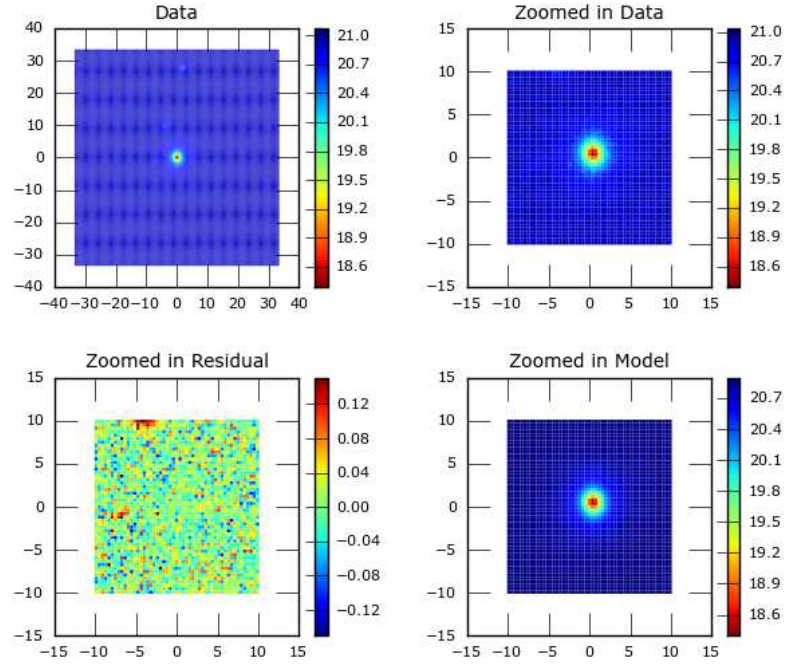
The object in Figure C2, B/T= 0.71 and  $R_d/R_b \sim 10$  is very likely to be an elliptical:  $p(\text{Ell})= 0.87$ . The Sersic and SerExp fits return almost the same magnitudes ( $M_r \sim -22.2$ ) and total half light radii ( $\sim 3.15''$ ). However,  $n = 7.15$  for the single Sersic fit, but  $n = 4.79$  for the SerExp bulge. For the SerExp, as for the deVExp fits, the second component is clearly necessary. The  $\chi^2_{dof}$  values for these fits are similar. In Figure C3, the Sersic magnitudes and half light radii are slightly larger, but otherwise the qualitative trends are the same: the single Sersic fit requires large  $n$ , and the second component in the SerExp fit clearly requires large  $R_d/R_b \sim 6$ .



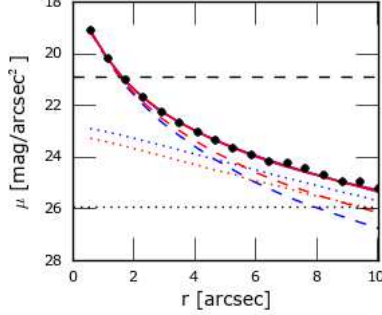
Note: The 1d data is calculated using background-subtracted data. The 2d data is shown with background included.

**Figure C2.** Example of an early-type galaxy with  $-21.5 < M_r < -22.5$ , large  $R_d/R_b$  and  $B/T \sim 0.7$ . Top left panel shows a  $\sim 20$  arcsec field centered on the object; top right panel provides a closer look. Middle right panel shows the best-fit SerExp model; middle left panel shows residuals from this fit. Bottom left panel shows the one-dimensional surface brightness profile (symbols), and our Sersic (solid magenta), deVExp (solid blue) and SerExp (solid red) fits; dotted and dashed curves show the corresponding disk and bulge components. Bottom right panel shows the associated residuals. Legend along the left shows the values of many quantities returned by the fits, and other information, such as the BAC  $p(\text{type})$ , for the object.

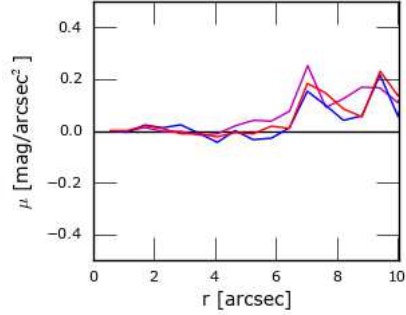
$z = 0.12$   
 $P(\text{Ell}) = 0.87$   
 $P(\text{S0}) = 0.10$   
 $P(\text{Sab}) = 0.02$   
 $P(\text{Scd}) = 0.02$   
 $M_{\text{ser}} = -22.373$   
 $M_{\text{serexp}} = -22.215$   
 $m_{\text{ser}} = 16.413$   
 $m_{\text{serexp}} = 16.571$   
 $B/T_{\text{serexp}} = 0.79$   
 $n_{\text{ser}} = 6.69$   
 $n_{\text{serexp}} = 4.80$   
 $r_{\text{hl,ser}} = 2.83$   
 $r_{\text{hl,serexp}} = 2.09$   
 $r_{\text{bulge,serexp}} = 1.38$   
 $r_{\text{disk,serexp}} = 8.12$



1D Sky Subtracted Surface Brightness Profile



1D Residual



Note: The 1d data is calculated using background-subtracted data. The 2d data is shown with background included.

**Figure C3.** Same as previous figure, but for another early-type galaxy selected at random from among those with the same  $M_r$  and  $B/T$  range.